Computer Networks –II

Paper Code MCA- 405

Contents

Written by

Parvinder Singh

Lecturer (Computer Science & Engineering)

Guru Jambheshwar University of Science & Technology, Hisar (India)

http://parvinder.50megs.com

Lessions 1, 3, 4, 6,7 and 8 vetted by

Dinesh Kumar, Reader (Computer Science & Engineering)

Guru Jambheshwar University of Science & Technology, Hisar (India)

Lession 2 and 5 vetted by

Dr. Sudhir Batra, Assistant Professor (Mathematics and Computer Science)

Technological Institute of Textile & Sciences, Bhiwani (India)

Paper Code: MCA- 405	**Author: Parvinder Singh**
Paper Name: Computer Networks-II	**Vetter: Dinesh Kumar**
Lesson Number: 01	

Introduction

Structure

1.9 References

1.0 Objective

This chapter introduces the concepts of interconnecting the different networks. The Internet protocol, which is used for addressing on internet, is also explained. The organization of different classes of IP address and mapping the internet address to physical address are also explained

1.1 Internetworking Concepts

In this section we will consider how we can assemble several smaller networks into larger ones and thus build towards the Internet. Lets us note two simple observations

- No single network can serve all users
- Users desire universal interconnection - access to global knowledge

These facts have driven the development of the Internet.

Both LAN and WAN are having technical limitations as-

- Local Area Networks which provide high speed communication are limited in geographic area.
- Wide Area Networks which serve large geographic area have speed limitations

No single network technology satisfies all needs.

1.1.1 Interconnection of Networks

Solution is to allow individual networks to rely on their own underlying technology and to establish a means by which they can communicate with one another. It is also desirable to have the network interconnectivity to appear transparent to the user.

1.1.2 Gateways

In order to connect individual networks (e.g LANs) to the outside world, computers called gateways are employed to provide all interconnections between physical networks.

Sometimes Gateways can also act as proxies (servers that give the false impression the internal network is directly connected to the outside world). The advantages of this scheme are

- Secure Mechanism
- Good for logging Internet activity.

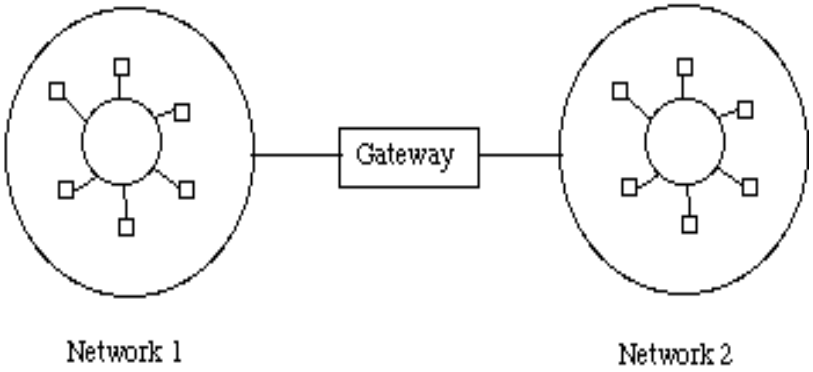


Figure 1.1 Gateway Connections

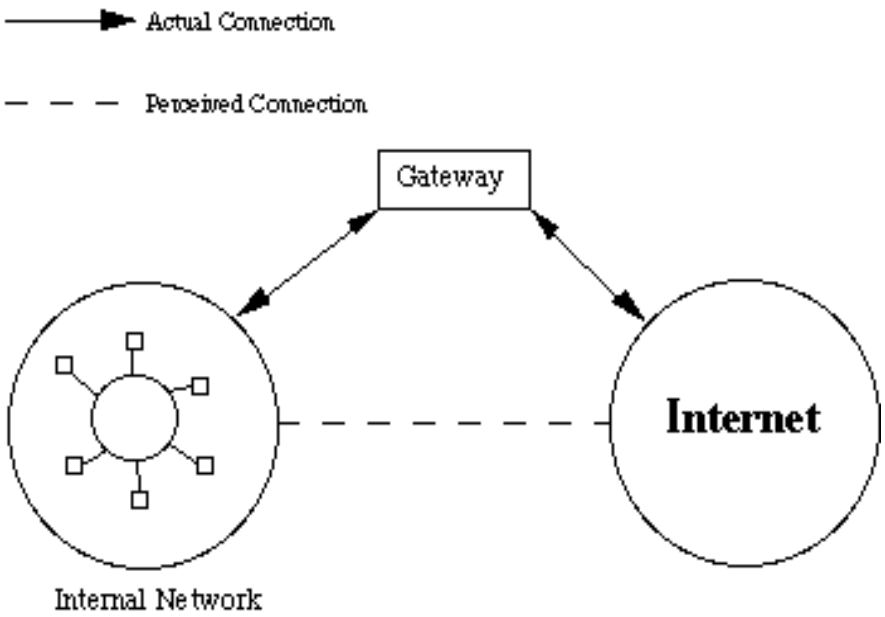


Figure 1.2 Proxy Gateway Connections

1.1.3 Routing

When one network is connected to another, a device called a router connects both networks and passes data between them. A router can be connected to more than one network. A router is, usually, a specialized computer with hardware and software specifically designed for routing network traffic. The route selected depends on traffic loads, what backbones are working etc. It should be noted that it is not necessary that all packets will be routed over the same route.

1.1.4 Some Routing Methods

Many different methods have been tried and many different strategies are in use today.

Routing Tables – These are specific maps (a list of routes) which tells that how to get somewhere. One by One all routes are tried from list until final destination is successfully reached.

- Centralised point – In this scheme all the traffic is sent through a centralized node in a network.
- Nearest Neighbor (Centralized adaptive routing) - Here a central node within each network knows only about its direct connections to the outside world. Each node sends traffic to nearest connection.

1.2 Internet Architecture

In 1989, Tim Berners Lee proposed a global hypertext project, to be known as the World Wide Web. It was designed to allow people to work together by combining their knowledge in a web of hypertext documents. Tim Berners Lee wrote the first World Wide Web server and the first client, a WYSIWYG (What You See Is What You Get) hypertext browser/editor. This work was started in October 1990, and the program “World Wide Web” was first made available within CERN in December, and on the Internet at large in the summer of 1991.

Through 1991 and 1993, Tim Berners Lee continued working on the design of the Web, coordinating feedback from users across the Internet. His initial specifications of URLs, HTTP and HTML were refined and used in larger circles as the web technology spread.

A browser, or viewer program is used to fetch and display “pages” of information from a server. A page is simply an ASCII text file, written using a simple meta language called Hypertext Markup Language (HTML).

1.2.1 Uniform Resource Locators - URLs

The URL is the basis of the WWW. We can think of a URL as an address that can lead us to any file on any machine anywhere in the world. Unlike the common postal address, however, these are written backwards. Actually backwards makes more sense. My postal adddress is:

Parvinder Singh
Deptt. of Computer Science & Engineering,
Guru Jambheshwar University of Science & Technology, Hisar,
Haryana

But if you want to deliver a letter to me, you first go to the Haryana, then Hisar, then Guru Jambhesahwar University of Science & Technology, then Deptt. of Computer Science & Engineering, then to Parvinder Singh. The URL is written in this more logical order.

A URL defines the location of a WWW page in the following way:

service:host:port/file and resource details.

For example:

http://www.gju.ernet.in/index.html

http://www.av.digital.com/cgi-bin/query?pg=q&what=web

It is not necessary that all URLs on the Web have to use the HTTP protocol. Some other URLs you might encounter are:

ftp (file transfer protocol)

news (for Usenet news groups)

telnet (for telnet)

mailto (to send email to a specific address)

1.2.2 Connection Establishment

To fetch a WWW page, the browser application process running on your local computer first establishes a connection to the remote host.

What this means is that the browser process uses the facilities of the network connecting the two computers to send a “connection request” message to a server process running on the computer whose name was given in the URL.

If the remote server process is prepared to accept the connection, it responds with a “connection accepted” message.

Note that we are, for the moment, ignoring the process of “looking up” the remote host - discovering the network address associated with its domain name.

1.2.3 HTTP Protocol

Once the two application processes have an established connection between them, they can communicate reliably. The browser then sends a request, in ordinary plain text, to the server, thus:

GET /home.html

The string GET something is one of many commands defined in the Hypertext Transfer Protocol, HTTP. The server responds by returning the contents of a file. Finally, the browser process interprets the HTML markup in the returned file, and displays it to the user.

1.3 IP Address Concept

Internet addresses can be symbolic or numeric. The symbolic form is easier to read, for example: myname@yahoo.com. The numeric form is a 32-bit unsigned binary value which is usually expressed in a dotted decimal format. For example, 9.167.5.8 is a valid Internet address. The numeric form is used by the IP

software. The mapping between the two is done by the Domain Name System. We shall first look at the numeric form, which is called the IP address.

The standards for IP addresses are described in RFC 1166 -- Internet Numbers.

To be able to identify a host on the internet, each host is assigned an address, the IP address, or Internet Address. When the host is attached to more than one network, it is called multi-homed and it has one IP address for each network interface. The IP address consists of a pair of numbers:

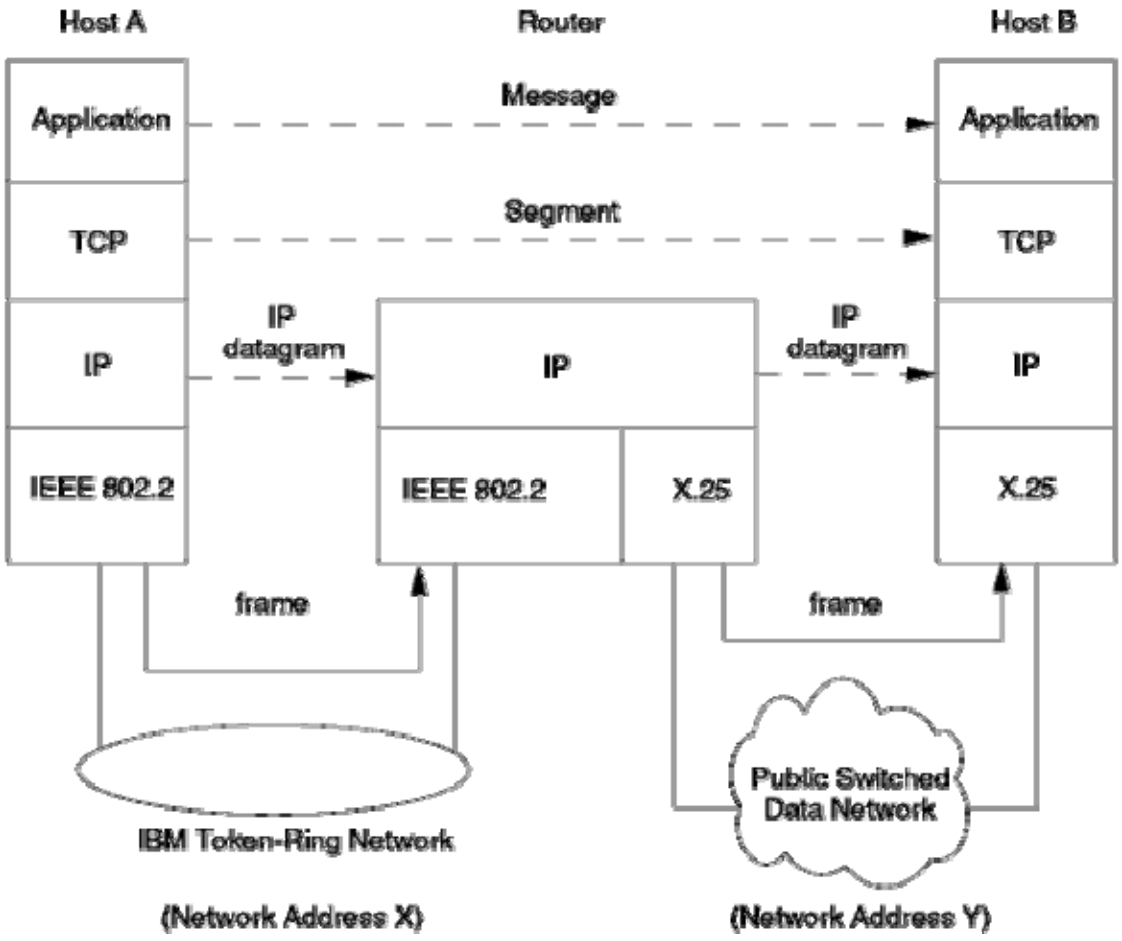


Figure 1.3 (IP Routing Mechanism)

IP address = <network number><host number>

The network number part of the IP address is centrally administered by the Internet Network Information Center (the InterNIC) and is unique throughout the Internet. (1)

IP addresses are 32-bit numbers usually represented in a dotted decimal form (as the decimal representation of four 8-bit values concatenated with dots). For example 128.2.7.9 is an IP address with 128.2 being the network number and 7.9 being the host number. The rules used to divide an IP address into its network and host parts are explained below.

The binary format of the IP address 128.2.7.9 is:

10000000 00000010 00000111 00001001

IP addresses are used by the IP protocol (see Internet Protocol (IP)) to uniquely identify a host on the internet. IP datagrams (the basic data packets exchanged between hosts) are transmitted by some physical network attached to the host and each IP datagram contains a source IP address and a destination IP address. To send a datagram to a certain IP destination, the target IP address must be translated or mapped to a physical address. This may require transmissions on the network to find out the destination's physical network address (for example, on LANs the Address Resolution Protocol, is used to translate IP addresses to physical MAC addresses).

The first bits of the IP address specify how the rest of the address should be separated into its network and host part.

The terms network address and netID are sometimes used instead of network number, but the formal term, used in RFC 1166, is network number. Similarly, the terms host address and hostID are sometimes used instead of host number.

There are five classes of IP addresses. These are shown in Figure 1.4

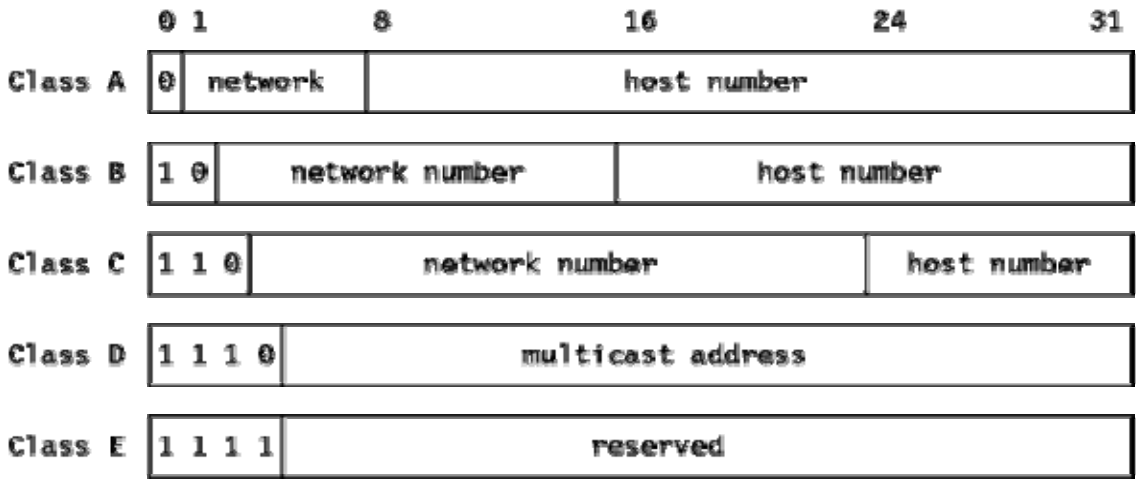


Figure 1.4

Note: Two numbers out of each of the class A, class B and class C network numbers, and two host numbers out of every network are pre-assigned: the “all bits 0” number and the “all bits 1” number. These are discussed below in Special IP Addresses.

Class A addresses use 7 bits for the network number giving 126 possible networks (we shall see below that out of every group of network and host numbers, two have a special meaning). The remaining 24 bits are used for the host number, so each networks can have up to $2^{24}-2$ (16,777,214) hosts.

Class B addresses use 14 bits for the network number, and 16 bits for the host number giving 16382 networks each with a maximum of 65534 hosts.

Class C addresses use 21 bits for the network number and 8 for the host number giving 2,097,150 networks each with up to 254 hosts.

Class D addresses are reserved for multicasting, which is used to address groups of hosts in a limited area. See Multicasting for more information on multicasting.

Class E addresses are reserved for future use.

It is clear that a class A address will only be assigned to networks with a huge number of hosts, and that class C addresses are suitable for networks with a small number of hosts. However, this means that medium-sized networks (those

with more than 254 hosts or where there is an expectation that there may be more than 254 hosts in the future) must use Class B addresses. The number of small- to medium-sized networks has been growing very rapidly in the last few years and it was feared that, if this growth had been allowed to continue unabated, all of the available Class B network addresses would have been used up till. This is termed the IP Address Exhaustion problem.

One point to note about the split of an IP address into two parts is that this split also splits the responsibility for selecting the IP address into two parts. The network number is assigned by the InterNIC, and the host number by the authority which controls the network. As we shall see in the next section, the host number can be further subdivided: this division is controlled by the authority which owns the network, and not by the InterNIC.

1.3.1 Subnets

Due to the explosive growth of the Internet, the use of assigned IP addresses became too inflexible to allow easy changes to local network configurations. These changes might occur when:

- A new physical network is installed at a location.
- Growth of the number of hosts requires splitting the local network into two or more separate networks.

To avoid having to request additional IP network addresses in these cases, the concept of subnets was introduced.

The host number part of the IP address is sub-divided again into a network number and a host number. This second network is termed a sub-network or subnet. The main network now consists of a number of subnets and the IP address is interpreted as:

<network number><subnet number><host number>

The combination of the subnet number and the host number is often termed the “local address” or the “local part”. “Sub-netting” is implemented in a way that is transparent to remote networks. A host within a network which has subnets is

aware of the sub-netting but a host in a different network is not; it still regards the local part of the IP address as a host number.

The division of the local part of the IP addresses into subnet number and host number parts can be chosen freely by the local administrator; any bits in the local part can be used to form the subnet accomplished. The division is done using a subnet mask which is a 32 bit number. Zero bits in the subnet mask indicate bit positions ascribed to the host number, and ones indicate bit positions ascribed to the subnet number. The bit positions in the subnet mask belonging to the network number are set to ones but are not used. Subnet masks are usually written in dotted decimal form, like IP addresses.

The special treatment of “all bits zero” and “all bits one” applies to each of the three parts of a sub-netted IP address just as it does to both parts of an IP address which has not been sub-netted. See Special IP Addresses. For example, a sub-netted Class B network, which has a 16-bit local part, could use one of the following schemes:

The first byte is the subnet number, the second the host number. This gives us 254 (256 minus 2 with the values 0 and 255 being reserved) possible subnets, each having up to 254 hosts. The subnet mask is 255.255.255.0.

The first 12 bits are used for the subnet number and the last four for the host number. This gives us 4094 possible subnets (4096 minus 2) but only 14 hosts per subnet (16 minus 2). The subnet mask is 255.255.255.240. There are many other possibilities.

While the administrator is completely free to assign the subnet part of the local address in any legal fashion, the objective is to assign a number of bits to the subnet number and the remainder to the local address. Therefore, it is normal to use a contiguous block of bits at the beginning of the local address part for the subnet number because this makes the addresses more readable (this is particularly true when the subnet occupies 8 or 16 bits). With this approach, either of the subnet masks above are “good” masks, but masks like 255.255.252.252 and 255.255.255.15 are not.

There are two types of subnetting: static and variable length. Variable length is the more flexible of the two. Which type of subnetting is available depends upon the routing protocol being used; native IP routing supports only static subnetting, as does the widely used RIP protocol. However, RIP Version 2 supports variable length subnetting as well.

1.3.1.1 Static Subnetting

Static subnetting means that all subnets in the subnetted network use the same subnet mask. This is simple to implement and easy to maintain, but it implies wasted address space for small networks. For example, a network of four hosts that uses a subnet mask of 255.255.255.0 wastes 250 IP addresses. It also makes the network more difficult to reorganize with a new subnet mask. Currently, almost every host and router supports static subnetting.

1.3.1.2 Variable Length Subnetting

When variable length subnetting is used, the subnets that make up the network may use different subnet masks. A small subnet with only a few hosts needs a subnet mask that accommodates only these few hosts. A subnet with many hosts attached may need a different subnet mask to accommodate the large number of hosts. The possibility to assign subnet masks according to the needs of the individual subnets will help conserve network addresses. Also, a subnet can be split into two parts by adding another bit to the subnet mask. Other subnets in the network are unaffected by the change. Not every host and router supports variable length subnetting.

Only networks of the size needed will be allocated and routing problems will be solved by isolating networks with routers that support variable subnetting. A host that does not support this kind of subnetting would have to route to a router that supports variable subnetting.

1.3.1.3 Mixing Static and Variable Length Subnetting

At first sight, it appears that the presence of a host which only supports static subnetting would prevent variable length subnetting from being used anywhere in the network. Fortunately this is not the case. Provided that the routers between subnets with different subnet masks are using variable length subnetting, the routing protocols employed are able to hide the difference between subnet masks from the hosts in a subnet. Hosts can continue to use basic IP routing and offload all of the complexities of the subnetting to dedicated routers.

1.3.2 Special IP Addresses

As noted above, any component of an IP address with a value “all bits 0” or all “all bits 1” has a special meaning.

All bits 0 stands for “this”: “this” host (IP address with <host number>=0) or “this” network (IP address with <network number>=0) and is only used when the real value is not known. This form is only used in source addresses when the host is trying to determine its IP addresses from a remote server. The host may include its host number if known, but not its subnet or network number.

All bits 1 stands for “all”: “all” networks or “all” hosts. For example, 128.2.255.255 (Classes B address with a host number of 255.255) means all hosts on network 128.2.

There is another address of special importance: the “all bits 1” class A network number 127 is reserved for the loop-back address. Anything sent to an address with 127 as the value of the high order byte, for example 127.0.0.1, must not be routed via a network but must be routed directly from the IP implementation's output driver to its input driver.

1.4 Address Resolution Protocol (ARP)

The address resolution protocol (ARP) is a protocol used by the Internet Protocol (IP) [RFC826], specifically IPv4, to map IP network addresses to the hardware addresses used by a data link protocol. The protocol operates below the network layer as a part of the interface between the OSI network and OSI link layer. It is used when IPv4 is used over Ethernet.

The term address resolution refers to the process of finding an address of a computer in a network. The address is “resolved” using a protocol in which a piece of information is sent by a client process executing on the local computer to a server process executing on a remote computer. The information received by the server allows the server to uniquely identify the network system for which the address was required and therefore to provide the required address. The address resolution procedure is completed when the client receives a response from the server containing the required address.

An Ethernet network uses two hardware addresses which identify the source and destination of each frame sent by the Ethernet. The destination address (all 1's) may also identify a broadcast packet (to be sent to all connected computers). The hardware address is also known as the Medium Access Control (MAC) address, in reference to the standards which define Ethernet. Each computer network interface card is allocated a globally unique 6 byte link address when the factory manufactures the card (stored in a PROM). This is the normal link source address used by an interface. A computer sends all packets which it creates with its own hardware source link address, and receives all packets which match the same hardware address in the destination field or one (or more) pre-selected broadcast/multicast addresses.

The Ethernet address is a link layer address and is dependent on the interface card which is used. IP operates at the network layer and is not concerned with the link addresses of individual nodes which are to be used. The address resolution protocol (ARP) is therefore used to translate between the two types of address. The ARP client and server processes operate on all computers using IP over Ethernet. The processes are normally implemented as part of the software driver that drives the network interface card.

There are four types of ARP messages that may be sent by the ARP protocol. These are identified by four values in the “operation” field of an ARP message. The types of message are:

1. ARP request
2. ARP reply

3. RARP request
4. RARP reply

The format of an ARP message is shown below:

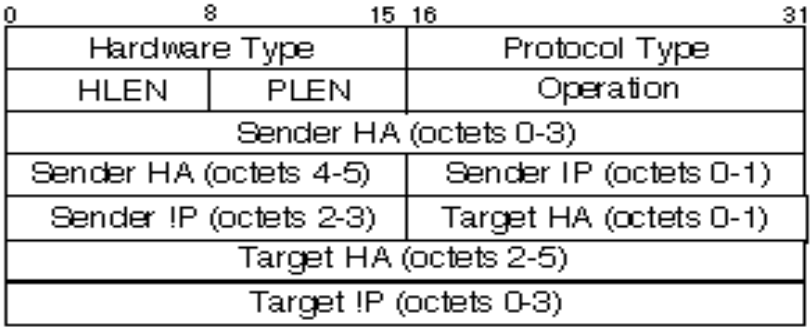


Figure 1.5 Format of an ARP message used to resolve the remote MAC Hardware Address (HA)

Here Hardware type is 16 bits and description of various values and their description is given below-

Value	Description
1	Ethernet.
2	Experimental Ethernet.
3	Amateur Radio AX.25.
4	Proteon ProNET Token Ring.
5	Chaos.
6	IEEE 802.
7	ARCNET.
8	Hyperchannel.
9	Lanstar.
10	Autonet Short Address.
11	LocalTalk.
12	LocalNet (IBM PCNet or SYTEK LocalNET).
13	Ultra link.
14	SMDS.
15	Frame Relay.
16	ATM, Asynchronous Transmission Mode.

17	HDLC.
18	Fibre Channel.
19	ATM, Asynchronous Transmission Mode.
20	Serial Line.
21	ATM, Asynchronous Transmission Mode.
22	MIL-STD-188-220.
23	Metricom.
24	IEEE 1394.1995.
25	MAPOS.
26	Twinaxial.
27	EUI-64.
28	HIPARP.
29	IP and ARP over ISO 7816-3.
30	ARPSec.
31	IPsec tunnel.
32	Infiniband.
33	CAI, TIA-102 Project 25 Common Air Interface.

Protocol type is also 16 bit and for IP value is 0x800.

HLEN- HLEN is 8 bit and it specifies the length of the hardware address in bytes.

PLEN- PLEN is 8 bit and it specifies the length of the protocol address in bytes.

Operation is opcode for various operations. It is 16 bit. The description of various values are given below-

Value	**Description**
1	Request.
2	Reply.
3	Request Reverse.
4	Reply Reverse.
5	DRARP Request.
6	DRARP Reply.
7	DRARP Error.
8	InARP Request.

9	InARP Reply.
10	ARP NAK.
11	MARS Request.
12	MARS Multi.
13	MARS MServ.
14	MARS Join.
15	MARS Leave.
16	MARS NAK.
17	MARS Unserv.
18	MARS SJoin.
19	MARS SLeave.
20	MARS Grouplist Request.
21	MARS Grouplist Reply.
22	MARS Redirect Map.
23	MAPOS UNARP.

To reduce the number of address resolution requests, a client normally caches resolved addresses for a (short) period of time. The ARP cache is of a finite size, and would become full of incomplete and obsolete entries for computers that are not in use if it was allowed to grow without check. The ARP cache is therefore periodically flushed of all entries. This deletes unused entries and frees space in the cache. It also removes any unsuccessful attempts to contact computers which are not currently running.

1.4.1 Example of use of the Address Resolution Protocol (ARP)

The figure below shows the use of ARP when a computer tries to contact a remote computer on the same LAN (known as “sysa”) using the "ping" program. It is assumed that no previous IP datagrams have been received form this computer, and therefore ARP must first be used to identify the MAC address of the remote computer.

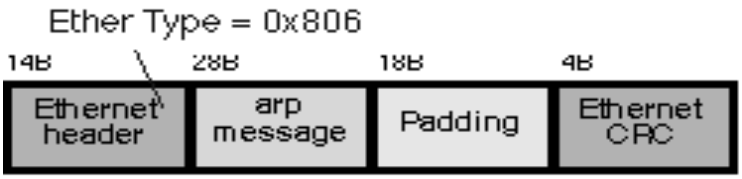


Figure 1.6

The ARP request message (“Who is X.X.X.X tell Y.Y.Y.Y”, where X.X.X.X and Y.Y.Y.Y are IP addresses) is sent using the Ethernet broadcast address, and an Ethernet protocol type of value 0x806. Since it is broadcast, it is received by all systems in the same collision domain (LAN). This is ensures that is the target of the query is connected to the network, it will receive a copy of the query. Only this system responds. The other systems discard the packet silently.

The target system forms an ARP response (“X.X.X.X is hh:hh:hh:hh:hh:hh”, where hh:hh:hh:hh:hh:hh is the Ethernet source address of the computer with the IP address of X.X.X.X). This packet is unicast to the address of the computer sending the query (in this case Y.Y.Y.Y). Since the original request also included the hardware address (Ethernet source address) of the requesting computer, this is already known, and doesn't require another ARP message to find this out.

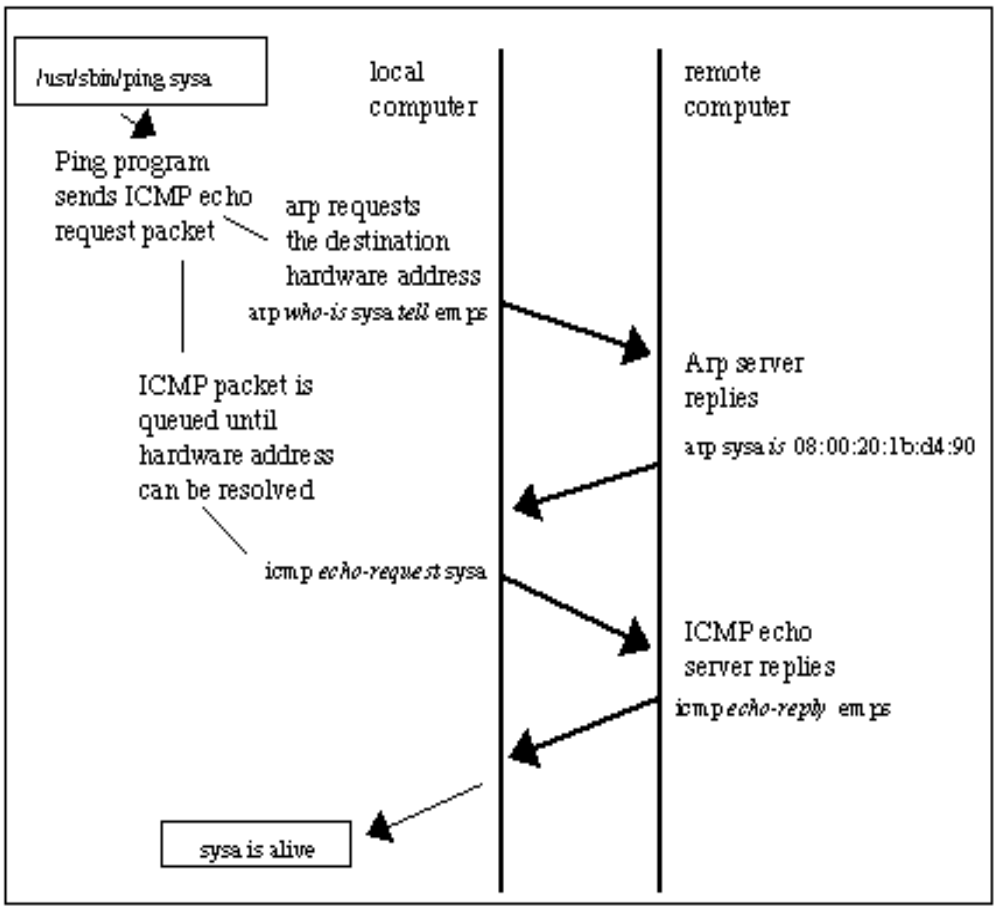


Figure 1.7

1.5 User Datagram Protocol

The User Datagram Protocol (UDP) is one of the core protocols of the Internet protocol suite. Using UDP, programs on networked computers can send short messages sometimes known as datagrams to one another.

UDP does not provide the reliability and ordering guarantees that TCP does. Datagrams may arrive out of order or go missing without notice. Without the overhead of checking if every packet actually arrived, UDP is faster and more efficient for many lightweight or time-sensitive purposes. Also, its stateless nature is useful for servers that answer small queries from huge numbers of clients. Compared to TCP, UDP is required for broadcast (send to all on local network) and multicast (send to all subscribers).

UDP is almost a null protocol; the only services it provides over IP are checksumming of data and multiplexing by port number. Therefore, an application program running over UDP must deal directly with end-to-end communication problems that a connection-oriented protocol would have handled - e.g., retransmission for reliable delivery, packetization and reassembly, flow control, congestion avoidance, etc., when these are required. The fairly complex coupling between IP and TCP will be mirrored in the coupling between UDP and many applications using UDP.

Common network applications that use UDP include the Domain Name System (DNS), streaming media applications, Voice over IP, Trivial File Transfer Protocol (TFTP) and online games.

The UDP header consists of only 4 fields of which two (Source Port and Checksum) are optional.

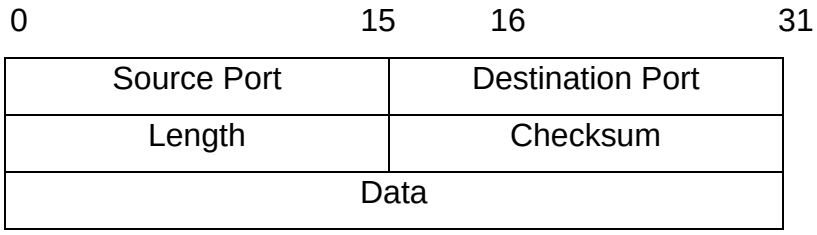


Figure 1.8 UDP header

Source Port (16 bits)- The port number of the sender. It is cleared to zero, if not used.

Destination port (16 bits)- This field identifies the destination port and is required.

Length (16 bits)- It specifies the length in bytes of the entire datagram: header and data. The minimum length is 8 bytes since that is the length of the header. The field size sets a theoretical limit of 65,527 bytes for the data carried by a single UDP datagram.

Checksum (16 bits)- This field is used for error-checking of the header and data. If the checksum is cleared to zero, then checksuming is disabled. If the computed checksum is zero, then this field must be set to 0xFFFF. Checksum is the 16-bit

one's complement of the one's complement sum of a pseudo header of information from the IP header, the UDP header, and the data, padded with zero octets at the end (if necessary) to make a multiple of two octets.

In other words, all 16-bit words are summed together using one's complement (with the checksum field set to zero). The sum is then one's complemented. This final value is then inserted as the checksum field. Algorithmically speaking, this is the same as for IPv4.

Data (Variable length)- This is the actual data transmitted.

When transported by IPv4, the pseudo header contains the following fields:

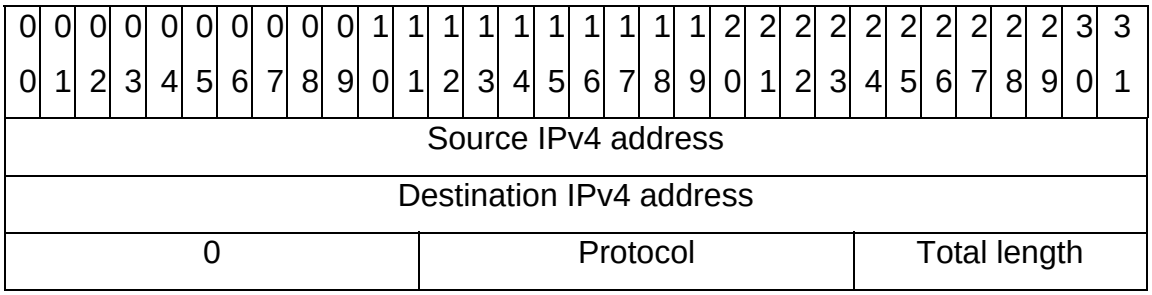


Figure 1.9 IPv4 Header

The checksum is not optional when transported by IPv6. In this case, the pseudo header contains the following fields:

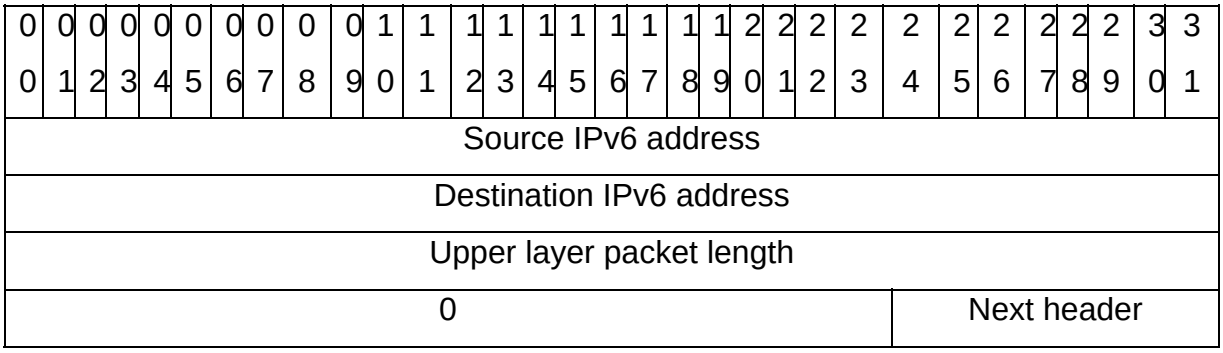


Figure 1.10 IPv6 Header

1.5.1 Ports

UDP utilizes ports to allow application-to-application communication. The port field is 16-bits so the valid range is 0 to 65,535. Port 0 is reserved, but is a

permissible source port value if the sending process does not expect messages in response. More details on port numbers is given in chapter 3.

1.6 Summary

As no single network can satisfy all users, the internetworking is required. The use of different technologies, hardware, software and different requirements from network make the internetworking a complex task. Further unique addressing on internet is provided by IP. As IP addresses are numeric, while user is more comfortable with symbolic addresses, so mapping between two is provided by DNS. ARP is a protocol used by the IP, to map IP network addresses to the hardware addresses used by a data link protocol. UDP does not provide the reliability but still make many applications such as voice chatting or video chatting possible.

1.7 Keywords

CERN- The acronym originally stood, in French, for Conseil Européen pour la Recherche Nucléaire (European Council for Nuclear Research), which was a provisional council for setting up the laboratory, established by 11 European governments in 1952. The acronym was retained for the new laboratory after the provisional council was dissolved, even though the name changed to the current Organisation Européenne pour la Recherche Nucléaire (European Organization for Nuclear Research) in 1954.

DNS- On the Internet, the Domain Name Service (DNS) stores and associates many types of information with domain names; most importantly, it translates domain names (computer hostnames) to IP addresses.

IPv6- Internet Protocol version 6 (IPv6) is a network layer protocol for packet-switched internetworks. It is designated as the successor of IPv4, the current version of the Internet Protocol, for general use on the Internet.

1.8 Self Assessment Questions

Q1. Explain concepts related to internetworking. Also explain routing and gateways

Q2. How IP addresses are classified? Explain different classes of IP addresses.

Q3. Explain the use of ARP.

Q4. In what situations UDP is used?

Q5 Write short notes on

i) Subnet ii) Special IP addresses iii) Ports iv) HTTP

Q6. Explain UDP header in detail.

Q7 Explain the format of ARP message.

Q8. What are the special IP addresses?

1.9 References

1. Computer Networks, Andrew S. Tanenbaum, Prentice Hall, 2002
2. Computer Networks, Randall Rustin, Pearson Education, 2000.
3. Computer Networks, V.S.Bagad, L.A.Dhotre, Technical Publications

Paper Code: MCA- 405	**Author: Parvinder Singh**
Paper Name: Computer Networks-II	**Vetter: Dr. Sudhir Batra**
Lesson Number: 02	

Internet Protocol

Structure

2.0 Objective

This chapter explains the concepts related to Internet protocol and Transmission Control Protocol. Concepts such as Fragmentation, IP Encapsulation, Datagram Forwarding, headers, acknowledgments and retransmissions of TCP are described in details.

2.1 Internet Protocol (IP)

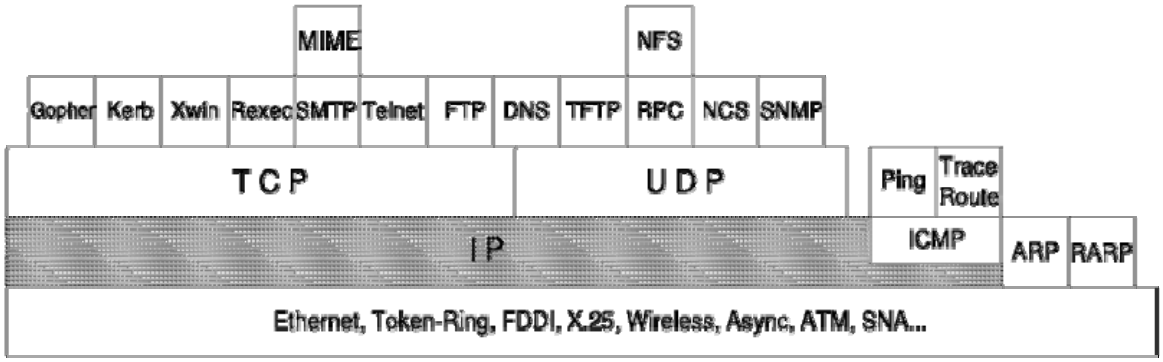


Figure 2.1 IP Protocol

IP is a standard protocol with STD number 5 which also includes ICMP (Internet Control Message Protocol) and IGMP (Internet Group Management Protocol). Its status is required.

The current IP specification can be found in RFCs 791, 950, 919 and 922, with updates in RFC 1349.

IP is a protocol that hides the underlying physical network by creating a virtual network view. It is an unreliable, best-effort connectionless packet delivery protocol.

It adds no reliability, flow control or error recovery to the underlying network interface protocol. Packets (datagrams) sent by IP may be lost, out of order, or even duplicated, and IP will not handle these situations. It is up to higher layers to provide these facilities.

IP also assumes little from the underlying network mechanisms, only that the datagrams will “probably” (best-effort) be transported to the addressed host.

2.1.1 IP Datagrams and Datagram Forwarding

The Internet datagram (IP datagram) is the base transfer packet in the Internet protocol suite. It has a header containing information for IP, and data that is relevant only to the higher level protocols. Figure 2.2 shows the base IP datagram.

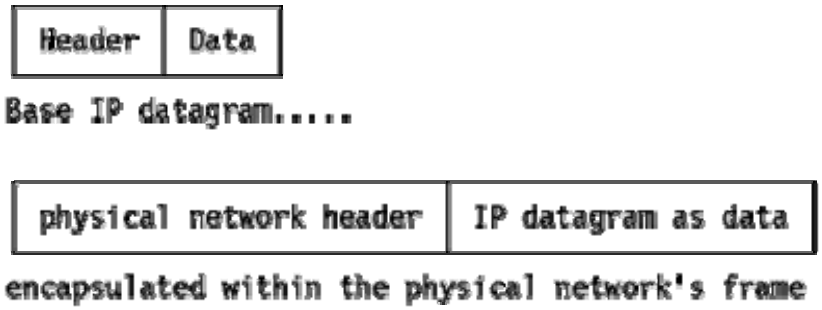


Figure 2.2

The IP datagram is encapsulated in the underlying network's frame, which usually has a maximum length or frame limitation, depending on the hardware used. For Ethernet, this will typically be 1500 bytes. Instead of limiting the IP datagram length to some maximum size, IP can deal with fragmentation and re-assembly of its datagrams. In particular, the IP standard does not impose a maximum size, but states that all sub-networks should be able to handle datagrams of at least 576 bytes.

Fragments of a datagram all have a header, basically copied from the original datagram, and data following it. They are treated as normal IP datagrams while being transported to their destination. Note, however, that if one of the fragments gets lost, the complete datagram is considered lost since IP does not provide any acknowledgment mechanism, so the remaining fragments will simply be discarded by the destination host.

2.1.2 IP Datagram Format

Figure 2.3 shows the format of IP Datagram. It is a minimum of 20 bytes long:

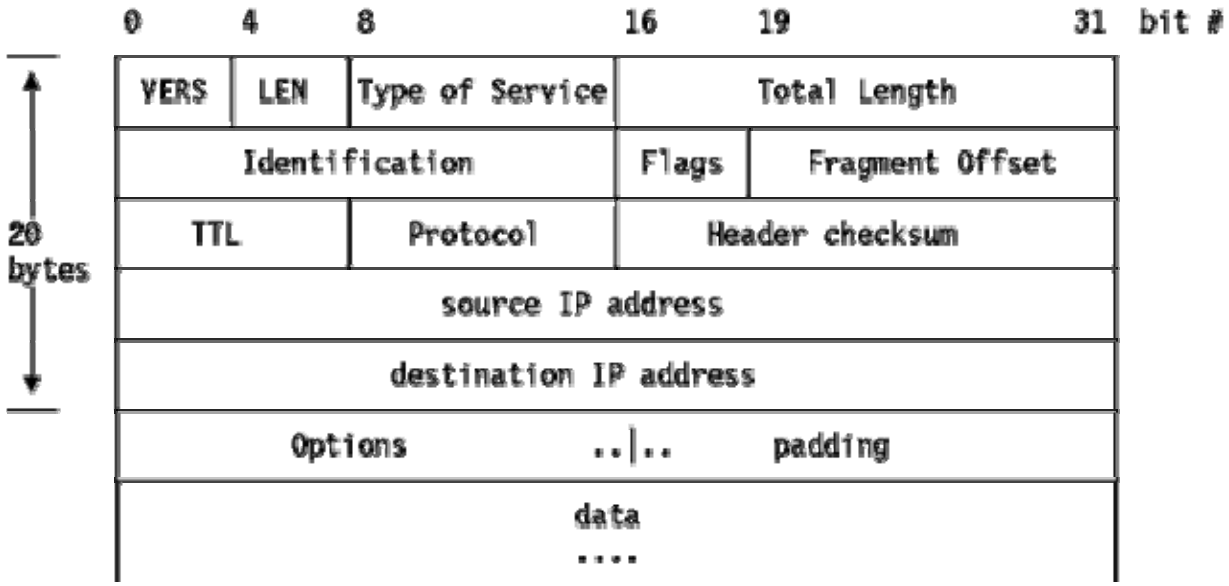


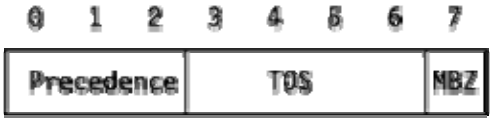
Figure 2.3 (Format of IP Datagram)

In the figure 2.3 :

VERS -The version of the IP protocol. Current is 4 and future is 6.

LEN - The length of the IP header counted in 32-bit quantities. This does not include the data field.

Type of Service - The type of service is an indication of the quality of service requested for this IP datagram.



Where:

Precedence is a measure of the nature and priority of this datagram:

000 - Routine

001 - Priority

010 - Immediate

011 - Flash

100 - Flash override

101 - Critical

110 – Inter-network control

111 - Network control

TOS Specifies the type of service value:

1000 - Minimize delay

0100 - Maximize throughput

0010 - Maximize reliability

0001 - Minimize monetary cost

0000 - Normal service

MBZ is reserved for future use ("must be zero" unless participating in an Internet protocol experiment which makes use of this bit).A detailed description of the type of service can be found in the RFC 1349.

Total Length - The total length of the datagram, header and data, specified in bytes.

Identification - A unique number assigned by the sender to aid in reassembling a fragmented datagram. Fragments of a datagram will have the same identification number.

Flags - Various control flags:

Where first cell is reserved, must be zero

DF - Don't Fragment: 0 means allow fragmentation, 1 means do not allow fragmentation.

MF - More Fragments: 0 means that this is the last fragment of this datagram, 1 means that this is not the last fragment.

Fragment Offset - Used with fragmented datagrams, to aid in reassembly of the full datagram. The value is the number of 64-bit pieces (header bytes are not counted) that are contained in earlier fragments. In the first (or only) fragment, this value is always zero.

Time to Live - Specifies the time (in seconds) this datagram is allowed to travel. Each router where this datagram passes is supposed to subtract from this field its processing time for this datagram. Actually a router is able to process a datagram in less than 1 second; thus it will subtract one from this field, and the TTL becomes a hop-count metric rather than a time metric. When the value reaches zero, it is assumed that this datagram has been traveling in a closed loop and it is discarded. The initial value should be set by the higher-level protocol which creates the datagram.

Protocol - Indicates the higher-level protocol to which IP should deliver the data in this datagram. Some important values are:

0 - Reserved

1- Internet Control Message Protocol (ICMP)

2 - Internet Group Management Protocol (IGMP)

3 - Gateway-to-Gateway Protocol (GGP)

4 - IP (IP encapsulation)

5 - Stream

6 - Transmission Control (TCP)

8 -Exterior Gateway Protocol (EGP)

9 - Private Interior Routing Protocol

17 - User Datagram (UDP)

89 -Open Shortest Path First

The full list can be found in STD 2 - Assigned Internet Numbers.

Header Checksum - Is a checksum on the header only. It does not include the data. The checksum is calculated as the 16-bit one's complement of the one's complement sum of all 16-bit words in the header. For the purpose of this calculation, the checksum field is assumed to be zero. If the header checksum does not match the contents, the datagram is discarded because at least one bit in the header is corrupt, and the datagram may even have arrived at the wrong destination.

Source IP Address - The 32-bit IP address of the host sending this datagram.

Destination IP Address - The 32-bit IP address of the destination host for this datagram.

Options - Variable length. An IP implementation is not required to be capable of generating options in the datagrams it creates, but all IP implementations are required to be able to process datagrams containing options. The Options field is variable in length. There may be zero or more options. There are two option formats. The format for each is dependent on the value of the option number found in the first byte.

Length - counts the length (in bytes) of the option, including the type and length fields.

Option data - contains data relevant to the option.

Padding - If an option is used, the datagram is padded with all-zero bytes up to the next 32-bit boundary.

Data - The data contained in the datagram is passed to a higher-level protocol, as specified in the protocol field.

2.1.3 IP Encapsulation

One of the most important concepts in inter-protocol operation is that of encapsulation. Most data originates within the higher layers of the OSI model. The protocols at these layers pass the data down to lower layers for transmission, usually in the form of discrete messages. Upon receipt, each lower-level protocol takes the entire contents of the message received and encapsulates it into its own message format, adding a header and possibly a footer that contain important control information. Encapsulation is explained in general terms in a separate topic.

A good analogy for how encapsulation works is a like sending of a letter enclosed in an envelope. You might write a letter and put it in a white envelope with a name and address, but if you gave it to a courier for overnight delivery, they would take that envelope and put it in a larger delivery envelope.

Due to the prominence of TCP/IP, the Internet Protocol is one of the most important places where data encapsulation occurs on a modern network. Data is passed to IP typically from one of the two main transport layer protocols: TCP or UDP. This data is already in the form of a TCP or UDP message with TCP or UDP headers. This is then encapsulated into the body of an IP message, usually called an IP datagram or IP packet. Encapsulation and formatting of an IP datagram is also sometimes called packaging—again, the implied comparison to an envelope is obvious.

2.1.4 Fragmentation

When an IP datagram travels from one host to another, it can cross different physical networks. Physical networks have a maximum frame size, called the Maximum Transmission Unit (MTU), which limits the length of a datagram that can be placed in one physical frame. Therefore, a scheme has been put in place to fragment long IP datagrams into smaller ones, and to reassemble them at the destination host. IP requires that each link has an MTU of at least 68 bytes, so if any network provides a lower value than this, fragmentation and re-assembly must be implemented in the network interface layer in a way that is transparent to IP. The sum of the maximum IP header length of 60 bytes and the minimum possible length of data in a non-final fragment (8 bytes) is 68. IP implementations are not required to handle unfragmented datagrams larger than 576 bytes, but most implementations will handle larger values, typically slightly more than 8192 bytes or higher, and rarely less than 1500.

An unfragmented datagram has all-zero fragmentation information. That is, the more fragments flag bit is zero and the fragment offset is zero. When fragmentation is to be done, the following steps are performed:

- The DF flag bit is checked to see if fragmentation is allowed. If the bit is set, the datagram will be discarded and an error will be returned to the originator using ICMP.

- Based on the MTU value, the data field is split into two or more parts. All newly created data portions must have a length which is a multiple of 8 bytes, with the exception of the last data portion.
- All data portions are placed in IP datagrams. The header of these datagrams are copies of the original one, with some modifications:
    - The more fragments flag bit is set in all fragments except the last.
    - The fragment offset field in each is set to the locate this data portion occupied in the original datagram, relative to the beginning of the original unfragmented datagram. The offset is measured in 8-byte units.
    - If options were included in the original datagram, the high order bit of the option type byte determines whether or not they will be copied to all fragment datagrams or just to the first one. For instance, source route options have to be copied in all fragments and therefore they have this bit set.
    - The header length field is of the new datagram is set.
    - The total length field of the new datagram is set.
    - The header checksum field is re-calculated.
- Each of these fragmented datagrams is now forwarded as a normal IP datagram. IP handles each fragment independently, that is, the fragments may traverse different routers to the intended destination, and they may be subject to further fragmentation if they pass through networks those have smaller MTUs.

Maximum Transmission Unit (MTU) refers to the size (in bytes) of the largest packet that a given layer of a communications protocol can pass onwards. MTU parameters usually appear in association with a communications interface.

At the destination host, the data has to be reassembled into one datagram. The identification field of the datagram was set by the sending host to a unique number (for the source host, within the limits imposed by the use of a 16-bit number). As fragmentation doesn't alter this field, incoming fragments at the receiving side can be identified, if this ID field is used together with the Source

and Destination IP addresses in the datagram. The Protocol field is also checked for this identification.

In order to reassemble the fragments, the receiving host allocates a buffer in storage as soon as the first fragment arrives. A timer routine is then started. When the timer timeouts and not all of the fragments have been received, the datagram is discarded. The initial value of this timer is called the IP datagram time-to-live (TTL) value. It is implementation dependent, and some implementations allow it to be configured; for example AIX Version 3.2 provides an ipfragttl option with a default value of 60 seconds.

When subsequent fragments of the datagram arrive, before the timer expires, the data is simply copied into the buffer storage, at the location indicated by the fragment offset field. As soon as all fragments have arrived, the complete original unfragmented datagram is restored, and the processing continues, just as for unfragmented datagrams.

The netstat command may be used on some TCP/IP hosts to list details of fragmentation that is occurring. An example of this is the netstat -i command in TCP/IP for OS/2.

2.2 Transmission Control Protocol (TCP)

TCP provides a virtual circuit (connection-oriented) communication service across the network. TCP includes rules for formatting messages, establishing and terminating virtual circuits, sequencing, flow control, and error correction. Most of the applications in the TCP/IP suite operate over the reliable transport service provided by TCP.

The TCP data unit is called a segment; the name is due to the fact that TCP does not recognize messages, but merely sends a block of bytes from the byte stream between sender and receiver. The fields of the segment (Figure 2.4) are:

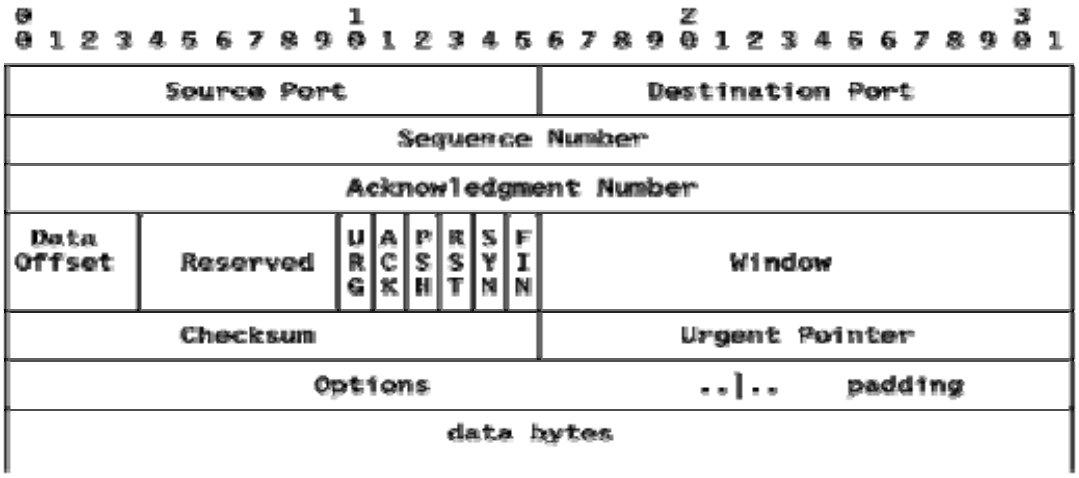


Figure2.4 (TCP Segment Format)

Source Port - The 16-bit source port number, used by the receiver to reply.

Destination Port - The 16-bit destination port number.

Sequence Number - The sequence number of the first data byte in this segment. If the SYN control bit is set, the sequence number is the initial sequence number (n) and the first data byte is n+1.

Acknowledgment Number - If the ACK control bit is set, this field contains the value of the next sequence number that the receiver is expecting to receive.

Data Offset - The number of 32-bit words in the TCP header. It indicates where the data begins.

Reserved - Six bits reserved for future use; must be zero.

URG - Indicates that the urgent pointer field is significant in this segment.

ACK - Indicates that the acknowledgment field is significant in this segment.

PSH - Push function.

RST - Resets the connection.

SYN - Synchronizes the sequence numbers.

FIN - No more data from sender.

Window - Used in ACK segments. It specifies the number of data bytes beginning with the one indicated in the acknowledgment number field which the receiver (= the sender of this segment) is willing to accept.

Checksum - The 16-bit one's complement of the one's complement sum of all 16-bit words in a pseudo-header, the TCP header and the TCP data. While computing the checksum, the checksum field itself is considered zero.

The pseudo-header is the same as that used by UDP for calculating the checksum. It is a pseudo-IP-header, only used for the checksum calculation, with the format shown in Figure 2.5:

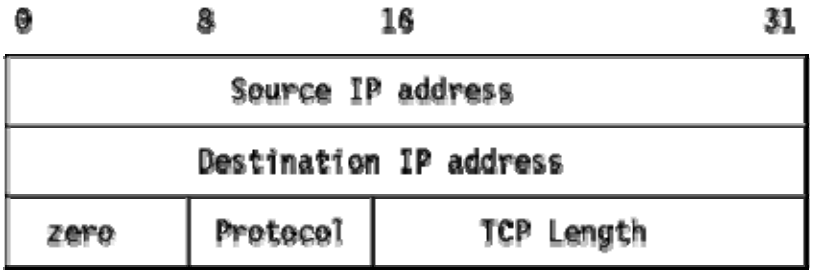


Figure 2.5 (Pseudo IP Header)

Urgent Pointer - Points to the first data octet following the urgent data. Only significant when the URG control bit is set.

Options - Just as in the case of IP datagram options, options can be either:

- A single byte containing the option number, or
- A variable length option in the following format:

Figure2.6 shows IP Datagram Option - Variable length option.

Figure 2.6 (IP Datagram Option)

There are currently only three options defined:

Kind	Length	Meaning
----	------	-------
0	-	End of option list.

1	-	No-Operation.
2	4	Maximum Segment Size.

Figure 2.7 shows Maximum Segment Size Option.

Figure 2.7 (Maximum Segment Size Option)

This option is only used during the establishment of the connection (SYN control bit set) and is sent from the side that is to receive data to indicate the maximum segment length it can handle. If this option is not used, any segment size is allowed.

Padding - All zero bytes used to fill up the TCP header to a total length that is a multiple of 32 bits.

2.2.1 Acknowledgments and Retransmissions

TCP sends data in variable length segments. Sequence numbers are based on a byte count. Acknowledgments specify the sequence number of the next byte that the receiver expects to receive.

Now suppose that a segment gets lost or corrupted. In this case, the receiver will acknowledge all further well-received segments with an acknowledgment referring to the first byte of the missing packet. The sender will stop transmitting when it has sent all the bytes in the window. Eventually, a timeout will occur and the missing segment will be retransmitted.

Suppose a window size of 1500 bytes, and segments of 500 bytes. Figure 19 describes the Acknowledgment and Retransmission Process.

A problem now arises, since the sender knows that segment 2 is lost or corrupted, but doesn't know anything about segments 3 and 4. The sender should at least retransmit segment 2, but it could also retransmit segments 3 and 4 (since they are within the current window). It is possible that:

Segment 3 has been received, and for segment 4 we don't know: it could be received, but ACK didn't reach us yet, or it could be lost also.

Segment 3 was lost, and we received the ACK 1500 upon the reception of segment 4.

Each TCP implementation is free to react to a timeout as the implementers wish. It could retransmit only segment 2, but in case 2 above, we will be waiting again until segment 3 times out. In this case, we lose all of the throughput advantages of the window mechanism. Or TCP might immediately resend all of the segments in the current window.

Whatever the choice, maximal throughput is lost. This is because the ACK does not contain a second acknowledgment sequence number indicating the actual frame received.

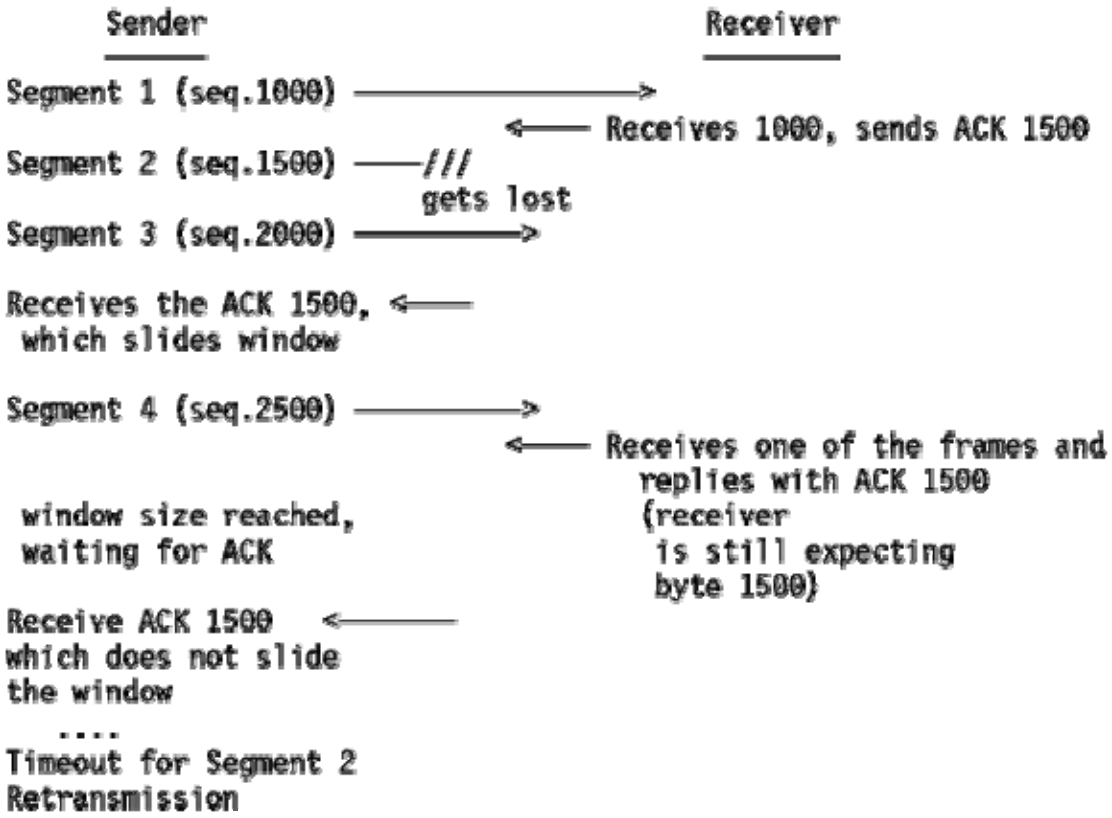


Figure 2.8 (Example of Acknowledgement and Retransmission)

2.2.2 Variable Timeout Intervals

Each TCP should implement an algorithm to adapt the timeout values to be used for the round trip time of the segments. To do this, TCP records the time at which a segment was sent, and the time at which the ACK is received. A weighted average is calculated over several of these round trip times, to be used as a timeout value for the next segment(s) to be sent.

This is an important feature, since delays may be variable on an internet, depending on multiple factors, such as the load of an intermediate low-speed network or the saturation of an intermediate IP gateway.

2.2.3 Establishing a TCP Connection

Before any data can be transferred, a connection has to be established between the two processes. One of the processes (usually the server) issues a passive OPEN call, the other an active OPEN call. The passive OPEN call remains dormant until another process tries to connect to it by an active OPEN.

On the network, three TCP segments are exchanged:

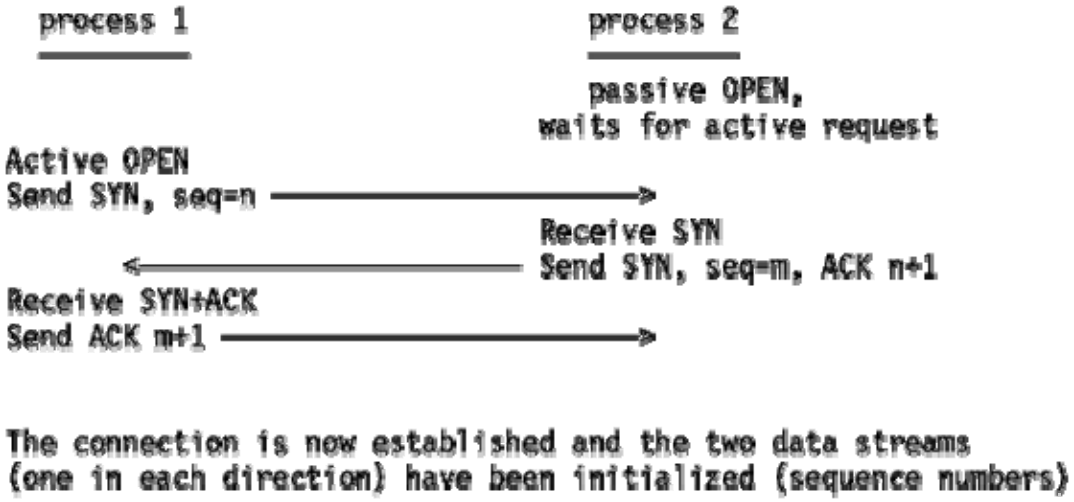


Figure 2.9 (Establishing a TCP Connection)

This whole process is known as **three-way handshake**. Note that the exchanged TCP segments include the initial sequence numbers from both sides, to be used on subsequent data transfers.

Closing the connection is done implicitly by sending a TCP segment with the FIN bit (no more data) set. As the connection is full-duplex (that is, we have two independent data streams, one in each direction), the FIN segment only closes the data transfer in one direction. The other process will now send the remaining data it still has to transmit and also ends with a TCP segment where the FIN bit is set. The connection is deleted (status information on both sides) once the data stream is closed in both directions.

2.2.4 TCP Segments Carried by IP Datagrams

TCP segments are transported in IP datagrams with the following parameter settings:

Type of Service = 00000000
that is: Precedence = routine
Delay = normal
Throughput = normal

Time to Live = 00111100 (1 minute)

2.3 Summary

IP does not provide the reassembly timer. It will treat each and every datagram, fragmented or not, the same way, that is, as individual messages. It is up to the higher layer to implement a timeout and to look after any missing fragments. The higher layer could be TCP for a connection-oriented transport network or the application for connectionless transport networks based upon UDP and IP.

2.4 Keywords

ICMP- The Internet Control Message Protocol (ICMP) is one of the core protocols of the Internet protocol suite. It is chiefly used by networked computers'

operating systems to send error messages—indicating, for instance, that a requested service is not available or that a host or router could not be reached.

IGMP- The Internet Group Management Protocol is a communications protocol used to manage the membership of Internet Protocol multicast groups. IGMP is used by IP hosts and adjacent multicast routers to establish multicast group memberships. It is an integral part of the IP multicast specification, like ICMP for unicast connections. IGMP can be used for online video and gaming, and allows more efficient use of resources when supporting these uses. IGMP does allow some attacks and firewalls commonly allow the user to disable it if it will not be needed.

Virtual Circuit- A virtual circuit (VC) is a communications arrangement in which data from a source user may be passed to a destination user over more than one real communications circuit during a single period of communication, but the switching is hidden from the users.

Time to Leave- Time to live (sometimes abbreviated TTL) is a limit on the period of time or number of iterations or transmissions in computer and computer network technology that a unit of data (e.g. a record) can experience before it should be discarded.

Handshaking- Handshaking is an automated process of negotiation that dynamically sets parameters of a communications channel established between two entities before normal communication over the channel begins. It follows the physical establishment of the channel and precedes normal information transfer.

EGP- The Exterior Gateway Protocol (EGP) is a (now 'obsolete') routing protocol for the Internet originally specified in 1982 by Eric C. Rosen of Bolt, Beranek and Newman, and David L. Mills.

2.5 Self Assessment Questions

Q1. Explain the format of IP datagram.

Q2. Explain steps of Fragmentation.

Q3. Explain the format of TCP segments.

Q4. How TCP connection is established?

Q5. Why IP encapsulation is important?

Q6. Explain Acknowledgments and Retransmissions in TCP.

2.6 References

1. Computer Network Architectures and Protocols, Paul Eliot Green, Plenum Publishing Company Limited, 1982
2. Schaum's Outline of Computer Networking, Ed Tittel, McGraw-Hill Professional, 2002
3. Data Networks, IP and the Internet: Networks, Protocols, Design, and Operation, Martin P. Clark, John Wiley and Sons, 2003

Paper Code: MCA- 405	**Author: Parvinder Singh**
Paper Name: Computer Networks-II	**Vetter: Dinesh Kumar**
Lesson Number: 03	

Routing Protocols

Structure

3.0 Objective

This chapter deals with routing and switching in detail. Different type of routing algorithms and the strategies employed in them are explained. Some popular routing protocols used in market are also described to understand the concepts practically.

3.1 What Is Routing?

Routing is the act of moving information across an internetwork from a source to a destination. Along the way, at least one intermediate node typically is encountered. Routing is often contrasted with bridging, which might seem to accomplish precisely the same thing to the casual observer. The primary difference between the two is that bridging occurs at Layer 2 (the link layer) of the OSI reference model, whereas routing occurs at Layer 3 (the network layer). This distinction provides routing and bridging with different information to use in the process of moving information from source to destination, so the two functions accomplish their tasks in different ways.

3.1.1 Routing Components

Routing involves two basic activities: determining optimal routing paths and transporting information groups (typically called packets) through an internetwork. In the context of the routing process, the latter of these is referred to as packet switching. Although packet switching is relatively straightforward, path determination can be very complex.

3.1.2 Path Determination

Routing protocols use metrics to evaluate what path will be the best for a packet to travel. A metric is a standard of measurement, such as path bandwidth, that is used by routing algorithms to determine the optimal path to a destination. To aid the process of path determination, routing algorithms initialize and maintain routing tables, which contain route information. Route information varies depending on the routing algorithm used.

Routing algorithms fill routing tables with a variety of information. Destination/next hop associations tell a router that a particular destination can be reached optimally by sending the packet to a particular router representing the “next hop” on the way to the final destination. When a router receives an incoming packet, it checks the destination address and attempts to associate this address with a next hop.

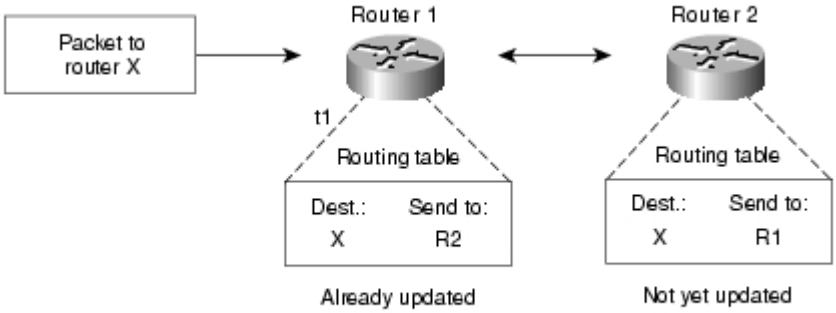


Figure 3.1 Destination/Next Hop Associations Determine the Data's Optimal Path

Figure 3.1 depicts a sample destination/next hop routing table.

Routing tables also can contain other information, such as data about the desirability of a path. Routers compare metrics to determine optimal routes, and these metrics differ depending on the design of the routing algorithm used. A variety of common metrics will be introduced and described later in this chapter.

Routers communicate with one another and maintain their routing tables through the transmission of a variety of messages. The routing update message is one such message that generally consists of all or a portion of a routing table. By analyzing routing updates from all other routers, a router can build a detailed picture of network topology. A link-state advertisement, another example of a message sent between routers, informs other routers of the state of the sender's links. Link information also can be used to build a complete picture of network topology to enable routers to determine optimal routes to network destinations.

3.1.3 Switching

Switching algorithms is relatively simple; it is the same for most routing protocols. In most cases, a host determines that it must send a packet to another host. Having acquired a router's address by some means, the source host sends a packet addressed specifically to a router's physical (Media Access Control [MAC]-layer) address, this time with the protocol (network layer) address of the destination host. As it examines the packet's destination protocol address, the router determines that it either knows or does not know how to forward the packet to the next hop. If the router does not know how to forward the packet, it typically drops the packet. If the router knows how to forward the packet, however, it changes the destination physical address to that of the next hop and transmits the packet.

The next hop may be the ultimate destination host. If not, the next hop is usually another router, which executes the same switching decision process. As the packet moves through the internetwork, its physical address changes, but its protocol address remains constant, as illustrated in Figure 3.2.

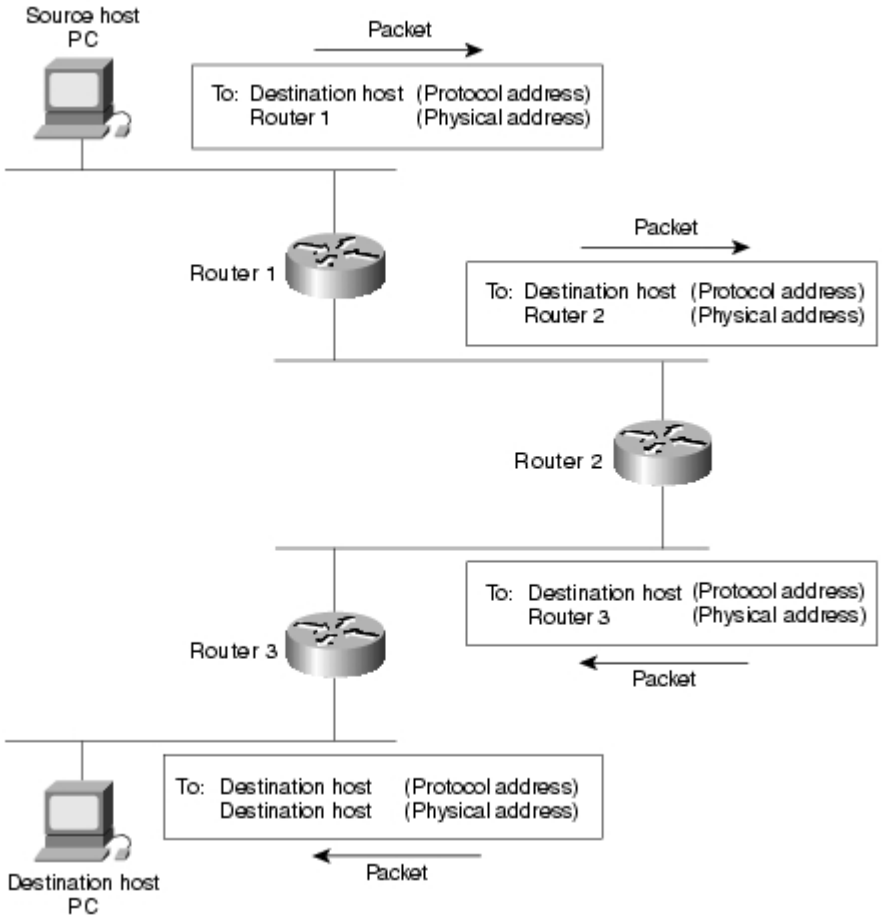


Figure 3.2 Switching Process

3.2 Routing Algorithms

Routing algorithms can be differentiated based on several key characteristics. First, the particular goals of the algorithm designer affect the operation of the resulting routing protocol. Second, various types of routing algorithms exist, and each algorithm has a different impact on network and router resources. Finally, routing algorithms use a variety of metrics that affect calculation of optimal routes. Routing algorithms can be classified by type. Key differentiators include these:

- Static versus dynamic
- Single-path versus multipath
- Flat versus hierarchical

• Host-intelligent versus router-intelligent
• Intradomain versus interdomain
• Link-state versus distance vector

3.2.1 Static Versus Dynamic

Static routing algorithms are hardly algorithms at all, but are table mappings established by the network administrator before the beginning of routing. These mappings do not change unless the network administrator alters them. Algorithms that use static routes are simple to design and work well in environments where network traffic is relatively predictable and where network design is relatively simple.

Because static routing systems cannot react to network changes, they generally are considered unsuitable for today's large, constantly changing networks. Most of the dominant routing algorithms today are dynamic routing algorithms, which adjust to changing network circumstances by analyzing incoming routing update messages. If the message indicates that a network change has occurred, the routing software recalculates routes and sends out new routing update messages. These messages permeate the network, stimulating routers to rerun their algorithms and change their routing tables accordingly.

Dynamic routing algorithms can be supplemented with static routes where appropriate. A router of last resort (a router to which all unroutable packets are sent), for example, can be designated to act as a repository for all unroutable packets, ensuring that all messages are at least handled in some way.

3.2.2 Single-Path Versus Multipath

Some sophisticated routing protocols support multiple paths to the same destination. Unlike single-path algorithms, these multipath algorithms permit traffic multiplexing over multiple lines. The advantages of multipath algorithms are obvious: They can provide substantially better throughput and reliability. This is generally called load sharing.

3.2.3 Flat Versus Hierarchical

Some routing algorithms operate in a flat space, while others use routing hierarchies. In a flat routing system, the routers are peers of all others. In a hierarchical routing system, some routers form what amounts to a routing backbone. Packets from nonbackbone routers travel to the backbone routers, where they are sent through the backbone until they reach the general area of the destination. At this point, they travel from the last backbone router through one or more nonbackbone routers to the final destination.

Routing systems often designate logical groups of nodes, called domains, autonomous systems, or areas. In hierarchical systems, some routers in a domain can communicate with routers in other domains, while others can communicate only with routers within their domain. In very large networks, additional hierarchical levels may exist, with routers at the highest hierarchical level forming the routing backbone.

The primary advantage of hierarchical routing is that it mimics the organization of most companies and therefore supports their traffic patterns well. Most network communication occurs within small company groups (domains). Because intradomain routers need to know only about other routers within their domain, their routing algorithms can be simplified, and, depending on the routing algorithm being used, routing update traffic can be reduced accordingly.

3.2.4 Host-Intelligent Versus Router-Intelligent

Some routing algorithms assume that the source end node will determine the entire route. This is usually referred to as source routing. In source-routing systems, routers merely act as store-and-forward devices, mindlessly sending the packet to the next stop.

Other algorithms assume that hosts know nothing about routes. In these algorithms, routers determine the path through the internetwork based on their own calculations. In the first system, the hosts have the routing intelligence. In the latter system, routers have the routing intelligence.

3.2.5 Intradomain Versus Interdomain

Some routing algorithms work only within domains; others work within and between domains. The nature of these two algorithm types is different. It stands to reason, therefore, that an optimal intradomain-routing algorithm would not necessarily be an optimal interdomain-routing algorithm.

3.2.6 Link-State Versus Distance Vector

Link-state algorithms (also known as shortest path first algorithms) flood routing information to all nodes in the internetwork. Each router, however, sends only the portion of the routing table that describes the state of its own links. In link-state algorithms, each router builds a picture of the entire network in its routing tables. Distance vector algorithms (also known as Bellman-Ford algorithms) call for each router to send all or some portion of its routing table, but only to its neighbors. In essence, link-state algorithms send small updates everywhere, while distance vector algorithms send larger updates only to neighboring routers. Distance vector algorithms know only about their neighbors.

Because they converge more quickly, link-state algorithms are somewhat less prone to routing loops than distance vector algorithms. On the other hand, link-state algorithms require more CPU power and memory than distance vector algorithms. Link-state algorithms, therefore, can be more expensive to implement and support. Link-state protocols are generally more scalable than distance vector protocols.

3.3 Routing Information Protocol

The Routing Information Protocol, or RIP, as it is more commonly called, is one of the most enduring of all routing protocols. RIP is also one of the more easily confused protocols because a variety of RIP-like routing protocols are in market, some of which even used the same name. RIP and the numerous RIP-like protocols were based on the same set of algorithms that use distance vectors to mathematically compare routes to identify the best path to any given destination

address. These algorithms emerged from academic research that dates back to 1957.

Today's open standard version of RIP, sometimes referred to as IP RIP, is formally defined in two documents: Request For Comments (RFC) 1058 and Internet Standard (STD) 56. As IP-based networks became both more numerous and greater in size, it became apparent to the Internet Engineering Task Force (IETF) that RIP needed to be updated. Consequently, the IETF released RFC 1388 in January 1993, which was then superceded in November 1994 by RFC 1723, which describes RIP 2 (the second version of RIP). These RFCs described an extension of RIP's capabilities but did not attempt to obsolete the previous version of RIP. RIP 2 enabled RIP messages to carry more information, which permitted the use of a simple authentication mechanism to secure table updates. More importantly, RIP 2 supported subnet masks, a critical feature that was not available in RIP.

We will here summarize the basic capabilities and features associated with RIP such as routing update process, RIP routing metrics, routing stability, and routing timers.

3.3.1 Routing Updates

RIP sends routing-update messages at regular intervals and when the network topology changes. When a router receives a routing update that includes changes to an entry, it updates its routing table to reflect the new route. The metric value for the path is increased by 1, and the sender is indicated as the next hop. RIP routers maintain only the best route (the route with the lowest metric value) to a destination. After updating its routing table, the router immediately begins transmitting routing updates to inform other network routers of the change. These updates are sent independently of the regularly scheduled updates that RIP routers send.

3.3.2 RIP Routing Metric

RIP uses a single routing metric (hop count) to measure the distance between the source and a destination network. Each hop in a path from source to destination is assigned a hop count value, which is typically 1. When a router receives a routing update that contains a new or changed destination network entry, the router adds 1 to the metric value indicated in the update and enters the network in the routing table. The IP address of the sender is used as the next hop.

3.3.3 RIP Stability Features

RIP prevents routing loops from continuing indefinitely by implementing a limit on the number of hops allowed in a path from the source to a destination. The maximum number of hops in a path is 15. If a router receives a routing update that contains a new or changed entry, and if increasing the metric value by 1 causes the metric to be infinity (that is, 16), the network destination is considered unreachable. The downside of this stability feature is that it limits the maximum diameter of a RIP network to less than 16 hops.

RIP includes a number of other stability features that are common to many routing protocols. These features are designed to provide stability despite potentially rapid changes in a network's topology. For example, RIP implements the split horizon and hold down mechanisms to prevent incorrect routing information from being propagated.

3.3.4 RIP Timers

RIP uses numerous timers to regulate its performance. These include a routing-update timer, a route-timeout timer, and a route-flush timer. The routing-update timer clocks the interval between periodic routing updates. Generally, it is set to 30 seconds, with a small random amount of time added whenever the timer is reset. This is done to help prevent congestion, which could result from all routers simultaneously attempting to update their neighbors. Each routing table entry has a route-timeout timer associated with it. When the route-timeout timer expires,

the route is marked invalid but is retained in the table until the route-flush timer expires.

3.3.5 RIP Packet Format

Figure 3.3 illustrates the IP RIP packet format.

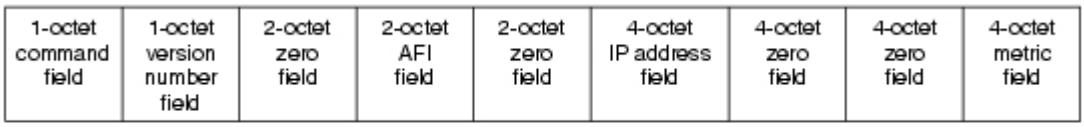


Figure 3.3 IP RIP Packet Format

The following descriptions summarize the IP RIP packet format fields illustrated in Figure 3.3:

Command—indicates whether the packet is a request or a response. The request asks that a router send all or part of its routing table. The response can be an unsolicited regular routing update or a reply to a request. Responses contain routing table entries. Multiple RIP packets are used to convey information from large routing tables.

Version number—specifies the RIP version used. This field can signal different potentially incompatible versions.

Zero—This field is not actually used by RFC 1058 RIP; it was added solely to provide backward compatibility with pre-standard varieties of RIP. Its name comes from its defaulted value: zero.

Address-family identifier (AFI)—Specifies the address family used. RIP is designed to carry routing information for several different protocols. Each entry has an address-family identifier to indicate the type of address being specified. The AFI for IP is 2.

Address—Specifies the IP address for the entry.

Metric—Indicates how many inter-network hops (routers) have been traversed in the trip to the destination. This value is between 1 and 15 for a valid route, or 16 for an unreachable route.

Please note that up to 25 occurrences of the AFI, Address, and Metric fields are permitted in a single IP RIP packet. (Up to 25 destinations can be listed in a single RIP packet.)

3.3.6 RIP 2 Packet Format

The RIP 2 specification (described in RFC 1723) allows more information to be included in RIP packets and provides a simple authentication mechanism that is not supported by RIP. Figure 3.4 shows the IP RIP 2 packet format.

Figure 3.4 IP RIP 2 Packet Format

The following descriptions summarize the IP RIP 2 packet format fields illustrated in Figure 3.4:

Command—indicates whether the packet is a request or a response. The request asks that a router send all or a part of its routing table. The response can be an unsolicited regular routing update or a reply to a request. Responses contain routing table entries. Multiple RIP packets are used to convey information from large routing tables.

Version—specifies the RIP version used. In a RIP packet implementing any of the RIP 2 fields or using authentication, this value is set to 2.

Unused—has a value set to zero.

Address-family identifier (AFI)—specifies the address family used. RIPv2's AFI field functions identically to RFC 1058 RIP's AFI field, with one exception: If the AFI for the first entry in the message is 0xFFFF, the remainder of the entry contains authentication information. Currently, the only authentication type is simple password.

Route tag—provides a method for distinguishing between internal routes (learned by RIP) and external routes (learned from other protocols).

IP address—specifies the IP address for the entry.

Subnet mask—contains the subnet mask for the entry. If this field is zero, no subnet mask has been specified for the entry.

Next hop—indicates the IP address of the next hop to which packets for the entry should be forwarded.

Metric—indicates how many internetwork hops (routers) have been traversed in the trip to the destination. This value is between 1 and 15 for a valid route, or 16 for an unreachable route.

Despite RIP's age and the emergence of more sophisticated routing protocols, it is far from obsolete. RIP is mature, stable, widely supported, and easy to configure. Its simplicity is well suited for use in stub networks and in small autonomous systems that do not have enough redundant paths to warrant the overheads of a more sophisticated protocol.

3.4 Open Shortest Path First (OSPF) Protocol

The Open Shortest Path First (OSPF) protocol is a link-state, hierarchical interior gateway protocol (IGP) for network routing. OSPF has two primary characteristics. The first is that the protocol is open, which means that its specification is in the public domain. The OSPF specification is published as Request For Comments (RFC) 1247. The second principal characteristic is that OSPF is based on the Shotest Path First (SPF) algorithm, which sometimes is referred to as the Dijkstra algorithm, named for the person credited with its creation. Suitable for complex networks with a large number of routers, OSPF provides equal cost multipath routing where packets to a single destination can be sent via more than one interface simultaneously. A link state database is constructed of the network topology which is identical on all routers in the area.

OSPF is perhaps the most widely used IGP in large networks. It can operate securely, using MD5 to authenticate peers before forming adjacencies, and before accepting link-state advertisements (LSA). A natural successor to the Routing Information Protocol (RIP), it was VLSM-capable or classless from its inception. A newer version of OSPF (OSPFv3) now supports IPv6 as well.

Multicast extensions to OSPF, the Multicast Open Shortest Path First (MOSPF) protocols, have been defined, but these are not widely used at present. OSPF can “tag” routes, and propagate the tags along with the routes.

An OSPF network can be broken up into smaller networks. A special area called the backbone area forms the core of the network, and other areas are connected to it. Inter-area routing goes via the backbone. All areas must connect to the backbone; if no direct connection is possible, a virtual link may be established.

Routers in the same broadcast domain or at each end of a point-to-point telecommunications link form adjacencies when they have detected each other. This detection occurs when a router “sees” itself in a hello packet. This is called a two way state and is the most basic relationship. The routers elect a designated router (DR) and a backup designated router (BDR) which act as a hub to reduce traffic between routers. OSPF uses both unicast and multicast to send “hello packets” and link state updates. Multicast addresses 224.0.0.5 and 224.0.0.6 are reserved for OSPF. In contrast to the Routing Information Protocol (RIP) or the Border Gateway Protocol (BGP), OSPF does not use TCP or UDP but uses IP directly.

3.4.1 Packet Format

All OSPF packets begin with a 24-byte header, as illustrated in Figure 3.5.

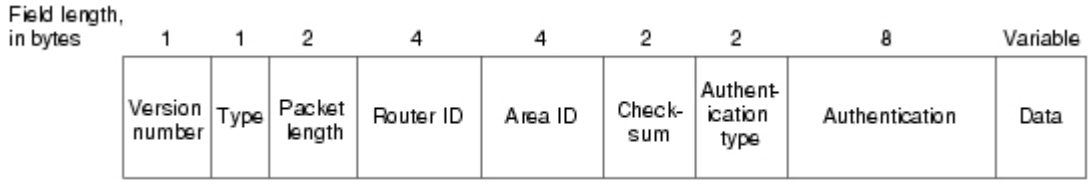


Figure 3.5 OSPF Packet

The following descriptions summarize the header fields illustrated in Figure 3.5.

Version number—identifies the OSPF version used.

Type—identifies the OSPF packet type as one of the following:

– Hello—establishes and maintains neighbor relationships.

– Database description—describes the contents of the topological database. These messages are exchanged when an adjacency is initialized.

– Link-state request—requests pieces of the topological database from neighbor routers. These messages are exchanged after a router discovers (by examining database-description packets) that parts of its topological database are outdated.

– Link-state update—responds to a link-state request packet. These messages also are used for the regular dispersal of LSAs. Several LSAs can be included within a single link-state update packet.

– Link-state acknowledgment—acknowledges link-state update packets.

Packet length—specifies the packet length, including the OSPF header, in bytes.

Router ID—identifies the source of the packet.

Area ID—identifies the area to which the packet belongs. All OSPF packets are associated with a single area.

Checksum—checks the entire packet contents for any damage suffered in transit.

Authentication type—contains the authentication type. All OSPF protocol exchanges are authenticated. The authentication type is configurable on per-area basis.

Authentication—contains authentication information.

Data—contains encapsulated upper-layer information.

3.5 Hello Protocol

The Open Shortest Path First (OSPF) routing protocol has a message type called Hello. The use of these messages is sometimes referred to as “The Hello Protocol”. OSPF is not directly related to the HELLO protocol described in this section, other than both protocols being used for routing in an Autonomous System(AS). It is possible OSPF borrowed the name Hello from the HELLO protocol.

The HELLO protocol is an interior protocol that uses a routing metric based on the length of time it takes a packet to make the trip between the source and the

destination. HELLO packets carry timestamp information which allows receivers to compute the shortest delay paths to destinations. The "best" route is the route with the shortest time delay. The unit of time used in HELLO is milliseconds. If a HELLO update packet takes less than 100 milliseconds to travel between two routers, a minimum value of 100 is used for that hop. Thus on networks built of high-speed interfaces, HELLO essentially defaults to using hop counts. As in any routing algorithm, HELLO cannot change routes too rapidly or it would be unstable. To avoid instabilities, implementations of HELLO build in hysteresis and "hesitate" to change routes until they have confidence that the change will be lasting.

The Hello Protocol is responsible for establishing and maintaining neighbour relationships. It also ensures that communication between neighbours is bidirectional. Hello packets are sent periodically out to all router interfaces. Bidirectional communication is indicated when the router sees itself listed in the neighbor's Hello Packet.

On multi-access networks, the Hello Protocol elects a Designated Router for the network. Among other things, the Designated Router controls what adjacencies will be formed over the network.

The Hello Protocol works differently on broadcast networks, as compared to non-broadcast networks. On broadcast networks, each router advertises itself by periodically multicasting Hello Packets. This allows neighbors to be discovered dynamically. These Hello Packets contain the router's view of the Designated Router's identity, and the list of routers whose Hello Packets have been seen recently.

On non-broadcast networks some configuration information is necessary for the operation of the Hello Protocol. Each router that may potentially become Designated Router has a list of all other routers attached to the network. A router, having Designated Router potential, sends Hello Packets to all other potential Designated Routers when its interface to the non- broadcast network first becomes operational. This is an attempt to find the Designated Router for the

network. If the router itself is elected Designated Router, it begins sending Hello Packets to all other routers attached to the network.

After a neighbor has been discovered, bidirectional communication ensured, and (if on a multi-access network) a Designated Router elected, a decision is made regarding whether or not an adjacency should be formed with the neighbor. An attempt is always made to establish adjacencies over point-to-point networks and virtual links. The first step in bringing up an adjacency is to synchronize the neighbors' topological databases.

3.5.1 Issues with Using Delay as a Metric

In theory, using delay calculations should result in more efficient route selection than simply using hop count, but at the cost of more complexity than a hop-count algorithm. This makes HELLO very interesting indeed, especially for a protocol that is over 20 years old. However, since the latency of a link is often unrelated to its bandwidth, using time delay as a link metric may lead to spurious results.

Furthermore, it is normal for the delay on any link to vary over time; in the case where there are two routes that are similar in cost, fluctuations in the delay for each route could result in rapid changes between routes (a phenomenon sometimes called route flapping). Adjustments are needed to the basic overview of the operation of the HELLO protocol above, to avoid these sorts of problems.

3.5.2 Current Role in TCP/IP

Like other early routing protocols, HELLO includes nothing fancy like authentication and so on; these features were not really needed in the early days of the Internet, when the internetworks were small and easily controlled. As the Internet grew, HELLO was eventually replaced by newer routing protocols such as RIP. It is now considered a obsolete protocol and is no longer used.

3.6 Mobile IP

As PDAs (Personal Digital Assistants) and the next generation of data-ready cellular phones become more widely deployed, a greater degree of connectivity

is almost becoming a necessity for the business user on the go. Data connectivity solutions for this group of users is a very different requirement than it is for the fixed dialup user or the stationary wired LAN user. Solutions here need to deal with the challenge of movement during a data session or conversation. Cellular service providers and network administrators wanting to deploy wireless LAN technologies need to have a solution which will grant this greater freedom.

Mobile IP is a tunneling-based solution. This tunneling enables a router on a user's home subnet to intercept and transparently forward IP packets to users while they roam beyond traditional network boundaries. This solution is a key enabler of wireless mobility, both in the wireless LAN arena, such as the 802.11 standard, and in the cellular environment for packet-based data offerings which offer connectivity to a user's home network and the Internet.

Mobile IP provides users the freedom to roam beyond their home subnet while consistently maintaining their home IP address. This enables transparent routing of IP datagrams to mobile users during their movement, so that data sessions can be initiated to them while they roam; it also enables sessions to be maintained in spite of physical movement between points of attachment to the Internet or other networks.

3.6.1 Benefits

Mobile IP is most useful in environments where mobility is desired and the traditional land line dial-in model or DHCP (Dynamic Host Configuration Protocol) do not provide adequate solutions for the needs of the users. If it is necessary or desirable for a user to maintain a single address while they transition between networks and network media, Mobile IP can provide them with this ability. Generally, Mobile IP is most useful in environments where a wireless technology is being utilized. This includes cellular environments as well as wireless LAN situations that may require roaming. Mobile IP can go hand in hand with many different cellular technologies like CDMA, TDMA, GSM, AMPS, as well as other proprietary solutions, to provide a mobile system which will scale for many users.

Each mobile node is always identified by its home address, no matter what its current point of attachment to the Internet, allowing for transparent mobility with respect to the network and all other devices. The only devices which need to be aware of the movement of this node are the mobile device and a router serving the user's topologically correct subnet.

3.6.2 Important terms in Mobile IP

Agent discovery—The method by which a mobile node determines whether it is currently connected to its home network or a foreign network and detects whether it has moved and the way it has moved. It is the mechanism by which mobile nodes query and discover mobility agents.

Care-of address—The termination point of the tunnel to a mobile node. This can be a collocated care-of address, where the mobile node acquires a local address and detunnels its own packets, or a foreign agent care-of address, where a foreign agent detunnels packets and forwards them to the mobile node.

Correspondent node—A peer with which a mobile node is communicating. A correspondent node may be either stationary or mobile.

Foreign agent—A router on a mobile node's visited network which provides routing services to the mobile node while registered. The foreign agent detunnels and delivers datagrams to the mobile node that were tunneled by the mobile node's home agent. For datagrams sent by a mobile node, the foreign agent may serve as a default router for registered mobile nodes.

Home address—An IP address that is assigned for an extended time to a mobile node. It remains unchanged regardless of where the node is attached to the Internet.

Home agent—A router on a mobile node's home network which tunnels packets to the mobile node while it is away from home. It keeps current location information for registered mobile nodes called a mobility binding.

Home network—The network or virtual network which matches the subnet address of the mobile node.

Mobile node—A host or router that changes its point of attachment from one network or subnet to another. A mobile node may change its location without changing its IP address; it may continue to communicate with other Internet nodes at any location using its home IP address, assuming link-layer connectivity to a point of attachment is available.

Mobility agent—A home agent or a foreign agent.

Mobility binding—The association of a home address with a care-of address and the remaining lifetime.

Mobility security association—A collection of security contexts between a pair of nodes, which may be applied to Mobile IP protocol messages exchanged between them. Each context indicates an authentication algorithm and mode, a secret (a shared key or appropriate public/private key pair), and a style of replay protection in use.

MTU—Maximum transmission unit. Maximum packet size, in bytes, that a particular interface can handle.

Node—A host or router.

Registration—The process by which the mobile node is associated with a care-of address on the home agent while it is away from home. This may happen directly from the mobile node to the home agent or through a foreign agent.

Security parameter index (SPI)—The index identifying a security context between a pair of nodes.

Tunnel—The path followed by a datagram while it is encapsulated from the home agent to the mobile node.

Virtual network—A network with no physical instantiation beyond a router (with a physical network interface on another network). The router (a home agent, for example) generally advertises reachability to the virtual network using conventional routing protocols.

Visited network—A network other than a mobile node's home network, to which the mobile node is currently connected.

Visitor list—The list of mobile nodes visiting a foreign agent.

3.6.3 Commands

Some important commands used in Mobile IP are given below-

Command	Purpose
router mobile	Enable Mobile IP on the router.
ip mobile home-agent	Enable home agent service.
ip mobile foreign-service	Enable foreign agent service on the interface.
show ip mobile globals	Check home agent and foreign agent global settings.
show ip route mobile	Check Mobile IP routes
show ip mobile traffic	Check protocol statistics.
no ip mobile home-agent	Disable home agent services.
no ip mobile foreign-agent	Disable foreign agent services.
no router mobile	Stop Mobile IP process.
show ip mobile host	Check mobile node counters (home agent only).
show ip mobile tunnel	Check active tunnels.
show ip mobile visitor	Check visitor bindings (foreign agent only).

3.7 Socket Interface

Both TCP and UDP make extensive use of the term "socket." A TCP socket represents the connection state between the local host and the remote peer. When talking about TCP connections which traverse the Internet, a socket is globally unique because it is described by 4 numbers: the local and remote IP addresses (32 bits each), and the local and remote port numbers (16 bits each).

Connections that do not traverse the Internet (e.g., between two hosts on an isolated LAN) are still unique within the attached network.

UDP sockets do not have the global uniqueness property, since they are not connection-oriented. For UDP, a socket really refers to just the local side.

For practical purposes, a socket is a structure in RAM that contains all the necessary state information. TCP sockets are considerably larger than UDP sockets since there is more connection state information to maintain. TCP sockets also require both a receive and a transmit buffer, whereas UDP sockets require only a receive buffer.

With Dynamic C version 6.57, each socket must have an associated tcp_Socket structure of 145 bytes or a udp_Socket structure of 62 bytes. The I/O buffers are in extended memory. For Dynamic C 7.30 these sizes are 136 bytes and 44 bytes, respectively.

For earlier versions of Dynamic C (than 6.57), each socket must have a tcp_Socket data structure that holds the socket state and I/O buffers. These structures are, by default, around 4200 bytes each. The majority of this space is used by the input and output buffers.

3.7.1 Port Numbers

Both TCP and UDP sockets make use of port numbers. Port numbers are a convenient method of allowing several simultaneous connections to exist between the same two hosts. Port numbers are also used to provide "well-known" starting points for common protocols. For example, TCP port number 23 is used for standard telnet connections. In general, port numbers below 1024 are used for standard services. Numbers between 1024 and 65535 are used for connections of a temporary nature. Often, the originator of a connection will select one of the temporary port numbers for its end of the connection, with the well-known number for the other end (which is often some sort of "server").

TCP and UDP port numbers are not related and operate in an independent "space." However, the well-known port numbers for TCP and UDP services often match if the same sort of protocol can be made to run over TCP or UDP.

When you open a socket using the TCP/IP libraries, you can specify a particular port number to use, or you can allow the library to pick a temporary port number for an "ephemeral" connection.

3.7.2 Opening TCP Sockets

There are two ways to open a TCP socket, passive and active. Passive open means that the socket is made available for connections originated from another host. This type of open is commonly used for Internet servers that listen on a well-known port, like 80 for HTTP (Hypertext Transfer Protocol) servers. Active open is used when the controller board is establishing a connection with another host which is (hopefully) listening on the specified port. This is typically used when the controller board is to be a "client" for some other server.

The distinction between passive and active open is lost as soon as the connection is fully established. When the connection is established, both hosts operate on a peer-to-peer basis. The distinction between who is "client" and who is "server" is entirely up to the application. TCP itself does not make a distinction.

3.7.3 Passive Open

To passively open a socket, call tcp_listen()or tcp_extlisten(); then wait for someone to contact your device. You supply the listen function with a pointer to a tcp_Socket data structure, the local port number others will be contacting on your device, and possibly the IP address and port number that will be acceptable for the peer. If you want to be able to accept connections from any IP address or any port number, set one or both to zero.

To handle multiple simultaneous connections, each new connection will require its own tcp_Socket and a separate call to one of the listen functions, but using the same local port number (lport value). The listen function will immediately return, and you must poll for the incoming connection. You can manually poll the socket using sock_established(). The proper procedure for fielding incoming connections is described below.

3.7.4 Active Open

When your Web browser retrieves a page, it actively opens one or more connections to the server's passively opened sockets. To actively open a connection, call tcp_open() or tcp_extopen(), which use parameters that are

similar to the ones used in the listen functions. Supply exact parameters for remip and port, which are the IP address and port number you want to connect to; the lport parameter can be zero, causing an unused local port between 1024 and 65535 to be selected.
If the open function returns zero, no connection was made. This could be due to routing difficulties, such as an inability to resolve the remote computer's hardware address with ARP. Even if non-zero is returned, the connection will not be immediately established. You will need to check the socket status.

3.7.5 Waiting for Connection Establishment

When you open a TCP socket either passively or actively, you must wait for a complete TCP connection to be established. This is technically known as the "3-way handshake." As the name implies, at least 3 packets must be exchanged between the peers. Only after completion of this process, which takes at least one round-trip time, does the connection become fully established such that application data transfer can proceed.

Unfortunately, the 3-way handshake may not always succeed: the network may get disconnected; the peer may cancel the connection; or the peer might even crash. The handshake may also complete, but the peer could immediately close or cancel the connection. These possibilities need to be correctly handled in a robust application. The consequences of not doing this right include locked-up sockets (i.e., inability to accept further connections) or protocol failures.

3.7.6 Opening and Closing a UDP Socket

udp_open() takes a remote IP address and a remote port number. If they are set to a specific value, all incoming and outgoing packets are filtered on that value (i.e., you talk only to the one remote address).

If the remote IP address is set to -1, the UDP socket receives packets from any valid remote address, and outgoing packets are broadcast. If the remote IP address is set to 0, no outgoing packets may be sent until a packet has been received. This first packet completes the socket, filling in the remote IP address and port number with the return address of the incoming packet. Multiple sockets can be opened on the same local port, with the remote address set to 0, to accept multiple incoming connections from separate remote hosts. When you are done communicating on a socket that was started with a 0 IP address, you can close it with sock_close() and reopen to make it ready for another source.

3.8 Summary

The most popular routing protocol is RIP. There are many flavors of this protocol available in market by different vendors. Latest in RIP series is RIP 2, it facilitates RIP messages to carry more information and use of a simple authentication mechanism to secure table updates. More importantly, RIP 2 supports subnet masks, a critical feature that was not available in RIP.

3.9 Keywords

DHCP- DHCP is a set of rules used by communications devices such as a computer, router or network adapter to allow the device to request and obtain an IP address from a server which has a list of addresses available for assignment.

PDA-Personal digital assistants (PDAs) are handheld computers that were originally designed as personal organizers, but became much more versatile over the years. PDAs are also known as pocket computers or palmtop computers. PDAs have many uses: calculation, use as a clock and calendar, playing computer games, accessing the Internet, sending and receiving E-mails, video recording, typewriting and word processing, use as an address book, making and writing on spreadsheets, use as a radio or stereo, and Global Positioning System (GPS). Newer PDAs also have both color screens and audio capabilities, enabling them to be used as mobile phones (smartphones), web browsers, or portable media players. Many PDAs can access the Internet, intranets or

extranets via Wi-Fi, or Wireless Wide-Area Networks (WWANs). One of the most significant PDA characteristic is the presence of a touch screen.

CDMA-Code division multiple access (CDMA) is a form of multiplexing and a method of multiple access that divides up a radio channel not by time (as in time division multiple access), nor by frequency (as in frequency-division multiple access), but instead by using different pseudo-random code sequences for each user.

TDMA-Time division multiple access (TDMA) is a channel access method for shared medium (usually radio) networks. It allows several users to share the same frequency channel by dividing the signal into different timeslots.

GSM-The Global System for Mobile Communications (GSM: originally from Groupe Spécial Mobile) is the most popular standard for mobile phones in the world. GSM service is used by over 2 billion people across more than 212 countries and territories.

AMPS-Advanced Mobile Phone System (AMPS) is the analog mobile phone system standard developed by Bell Labs, and officially introduced in the America in 1983 although its use has dropped considerably with the introduction of various digital standards.

3.10 Self Assessment Questions

Q1. What is routing? Explain different types of Routing Algorithms.

Q2. Why is RIP so popular? Explain RIP stability features.

Q3. Explain RIP 2 packet format in detail.

Q4. Explain OSPF protocol in detail

Q5 Write short notes on

i) Hello Protocol ii) Mobile IP iii) Socket and Ports iv) RIP Timer

Q6. Explain some important terms and commands in Mobile IP.

Q7. Why role of Hello Protocol is limited in today's environment ?

3.11 Reference

1. Schaum's Outline of Computer Networking, Ed Tittel, McGraw-Hill Professional, 2002

2. Data Networks, IP and the Internet: Networks, Protocols, Design, and Operation, Martin P. Clark, John Wiley and Sons, 2003

Paper Code: MCA- 405	**Author: Parvinder Singh**
Paper Name: Computer Networks-II	**Vetter: Dinesh Kumar**
Lesson Number: 04	

WAN AND LAN

Structure

4.0 Objective

In this lession different WAN and LAN technologies are described. Circuit Switching Networks and Packet Switching Networks are covered to explain the concepts related to WAN. Switches, Routers and Bridges are explained to have an idea of LAN devices.

4.1 LAN, WAN and MAN

The Local Area Network (LAN) is by far the most common type of data network. As the name suggests, a LAN serves a local area (typically the area of a floor of a building, but in some cases spanning a distance of several kilometers). Typical installations are in industrial plants, office buildings, college or university campuses, or similar locations. In these locations, it is feasible for the owning Organization to install high quality, high-speed communication links interconnecting nodes. Typical data transmission speeds are one to 100 megabits per second.

A Wide Area Network (WAN) is a data communications network that covers a relatively broad geographic area and that often uses transmission facilities provided by common carriers, such as telephone companies. WAN technologies generally function at the lower three layers of the OSI reference model: the physical layer, the data link layer, and the network layer.

A Metropolitan Area Network (MAN) is a relatively new class of network, it serves a role similar to an ISP, but for corporate users with large LANs. There are three important features which discriminate MANs from LANs or WANs:

- The network size falls intermediate between LANs and WANs. A MAN typically covers an area of between 5 and 50 km diameter. Many MANs cover an area the size of a city, although in some cases MANs may be as small as a group of buildings.
- A MAN (like a WAN) is not generally owned by a single organization. The MAN, its communications links and equipment are generally owned by either a consortium of users or by a single network provider who sells the service to the users. This level of service provided to each user must

therefore be negotiated with the MAN operator, and some performance guarantees are normally specified.

- A MAN often acts as a high speed network to allow sharing of regional resources (similar to a large LAN). It is also frequently used to provide a shared connection to other networks using a link to a WAN.

4.2 Point-to-Point Links

A point-to-point link provides a single, pre-established WAN communications path from the customer premises through a carrier network, such as a telephone company, to a remote network. Point-to-point lines are usually leased from a carrier and thus are often called leased lines. For a point-to-point line, the carrier allocates pairs of wire and facility hardware to your line only. These circuits are generally priced based on bandwidth required and distance between the two connected points. Point-to-point links are generally more expensive than shared services such as Frame Relay. Figure 4.1 illustrates a typical point-to-point link through a WAN.

Figure 4.1 Point-to-Point Link

4.3 Circuit Switching Networks

Switched circuits allow data connections that can be initiated when needed and terminated when communication is complete. This works much like a normal telephone line works for voice communication. Integrated Services Digital Network (ISDN) is a good example of circuit switching. When a router has data for a remote site, the switched circuit is initiated with the circuit number of the remote network. In the case of ISDN circuits, the device actually places a call to the telephone number of the remote ISDN circuit. When the two networks are

connected and authenticated, they can transfer data. When the data transmission is complete, the call can be terminated. Figure 4.2 illustrates an example of this type of circuit. In this figure DCE is Data Circuit- Terminating Equipment.

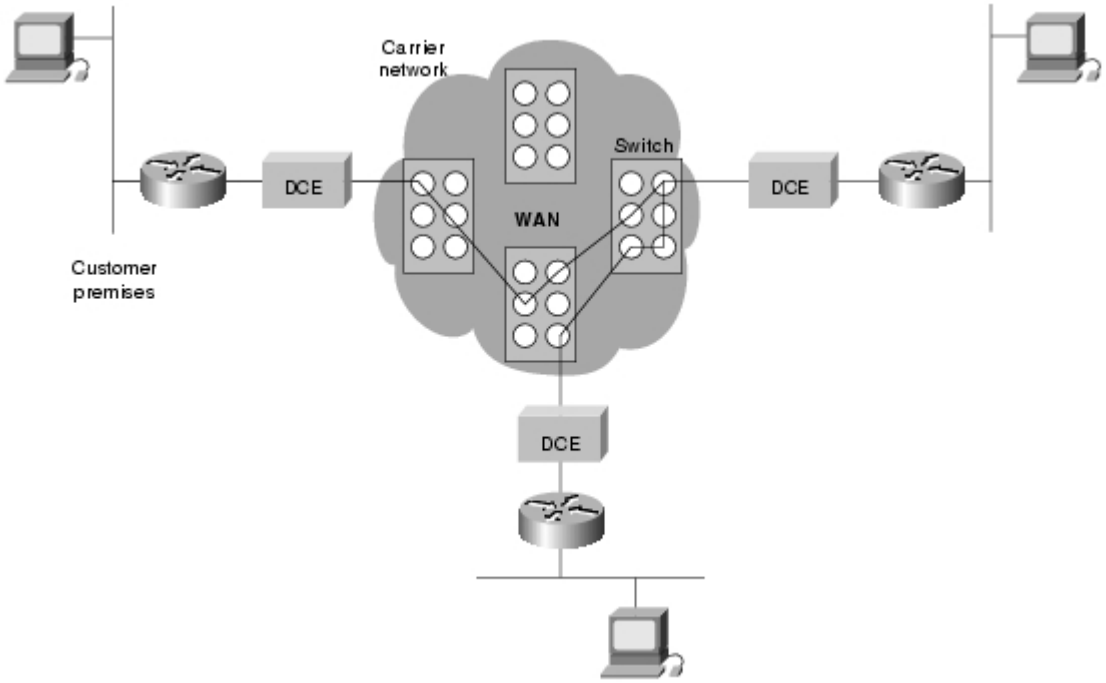


Figure 4.2 Circuit-Switched WAN

Each user has sole access to a circuit (functionally equivalent to a pair of copper wires) during network use. Consider communication between two points A and D in a network. The connection between A and D is provided using (shared) links between two other pieces of equipment, B and C.

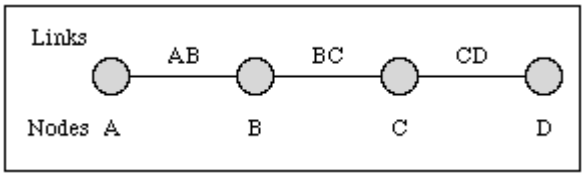


Figure 4.3 A connection between two systems A & D formed from 3 links

Network use is initiated by a connection phase, during which a circuit is set up between source and destination, and terminated by a disconnect phase. These phases, with associated timings, are illustrated in the figure below.

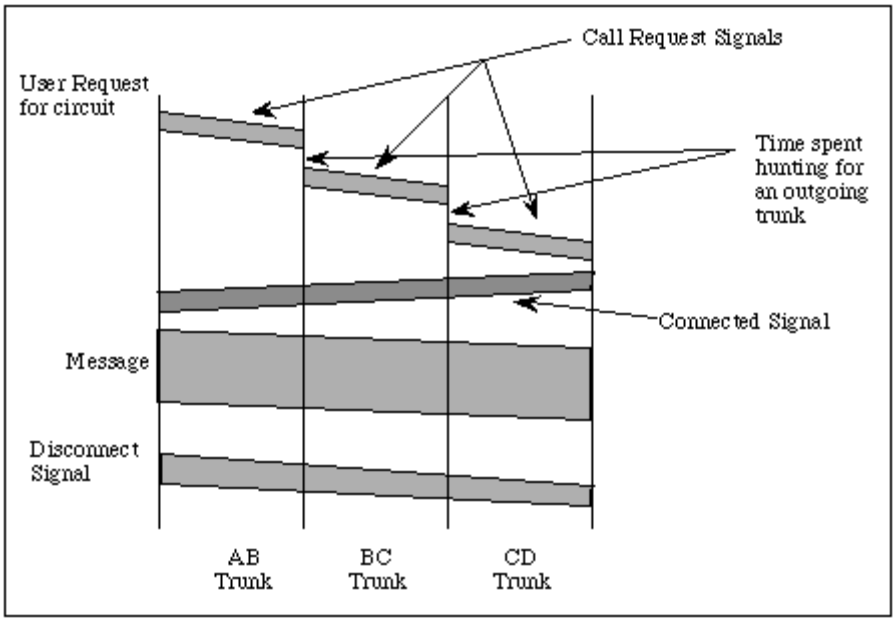


Figure 4.4 A circuit switched connection between A and D

After a user requests a circuit, the desired destination address must be communicated to the local switching node (B). In a telephony network, this is achieved by dialing the number.

Node B receives the connection request and identifies a path to the destination (D) via an intermediate node (C). This is followed by a circuit connection phase handled by the switching nodes and initiated by allocating a free circuit to C (link BC), followed by transmission of a call request signal from node B to node C. In turn, node C allocates a link (CD) and the request is then passed to node D after a similar delay.

The circuit is then established and may be used. While it is available for use, resources (i.e. in the intermediate equipment at B and C) and capacity on the links between the equipment are dedicated to the use of the circuit.

After completion of the connection, a signal confirming circuit establishment (a connect signal in the diagram) is returned; this flows directly back to node A with no search delays since the circuit has been established. Transfer of the data in the message then begins. After data transfer, the circuit is disconnected; a simple disconnect phase is included after the end of the data transmission. Delays for setting up a circuit connection can be high, especially if ordinary telephone equipment is used. Call setup time with conventional equipment is typically the order of 5 to 25 seconds after completion of dialing. New fast circuit switching techniques can reduce delays. Trade-offs between circuit switching and other types of switching depend strongly on switching times.

4.4 Packet Switching Networks

Packet switching Networks allow users to share common carrier resources. Because this allows the carrier to make more efficient use of its infrastructure, the cost to the customer is generally much better than with point-to-point lines. In a packet switching setup, networks have connections into the carrier's network, and many customers share the carrier's network. The carrier can then create virtual circuits between customers sites by which packets of data are delivered from one to the other through the network. The section of the carrier's network that is shared is often referred to as a cloud.

Some examples of packet-switching networks include Asynchronous Transfer Mode (ATM), Frame Relay, Switched Multimegabit Data Services (SMDS), and X.25. Figure 4.5 show an example packet-switched network.

The virtual connections between customer sites are often referred to as a virtual circuit.

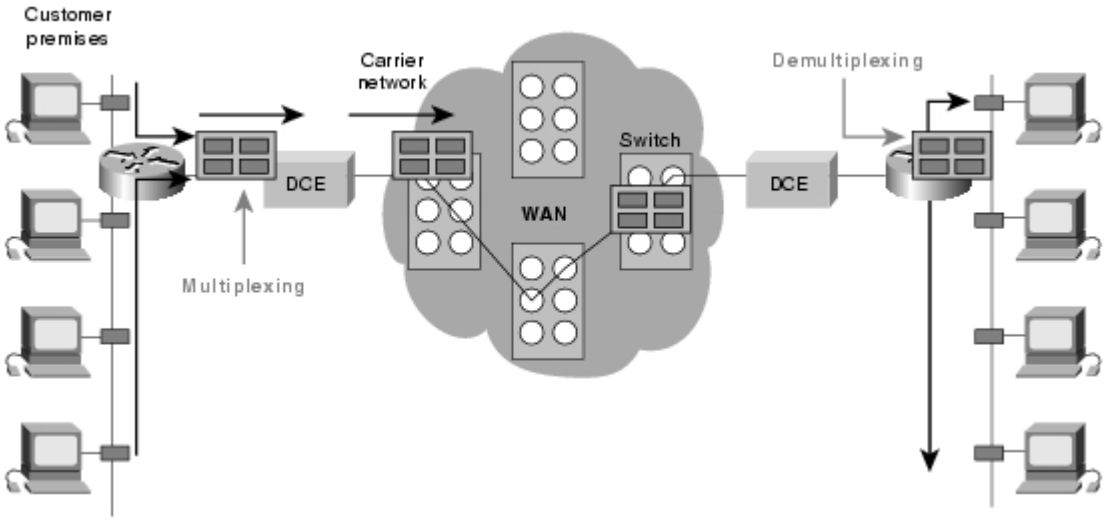


Figure 4.5 Packet Switching Network

The fundamental difference in packet communication is that the data is formed into packets with a pre-defined header format (i.e. PCI), and well-known "idle" patterns which are used to occupy the link when there is no data to be communicated.

A packet network equipment discards the "idle" patterns between packets and processes the entire packet as one piece of data. The equipment examines the packet header information (PCI) and then either removes the header (in an end system) or forwards the packet to another system. If the out-going link is not available, then the packet is placed in a queue until the link becomes free. A packet network is formed by links which connect packet network equipment.

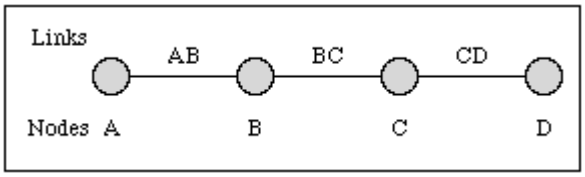


Figure 4.6 Communication between A and D using circuits which are shared using packet switching

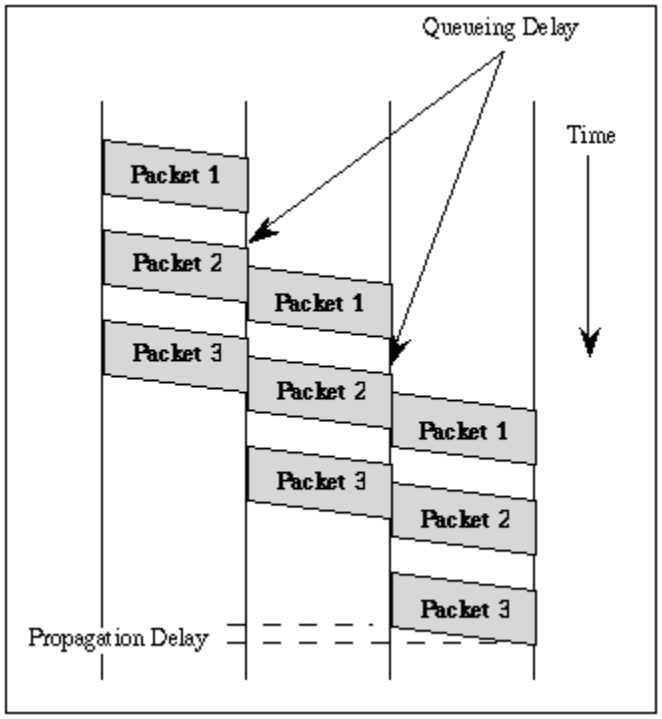


Figure 4.7 Packet-switched communication between systems A and D

(The message in this case has been broken into three parts labeled 1-3)

There are two important benefits from packet switching.

- The first and most important benefit is that since packets are short, the communication links between the nodes are only allocated to transferring a single message for a short period of time while transmitting each packet. Longer messages require a series of packets to be sent, but do not require the link to be dedicated between the transmission of each packet. The implication is that packets belonging to other messages may be sent between the packets of the message being sent from A to D. This provides a much fairer sharing of the resources of each of the links.
- Another benefit of packet switching is known as "pipelining". Pipelining is visible in the figure 4.7. At the time packet 1 is sent from B to C, packet 2 is sent from A to B; packet 1 is sent from C to D while packet 2 is sent from B to C, and packet 3 is sent from A to B, and so forth. This simultaneous use of communications links represents a gain in efficiency, the total delay for transmission across a packet network may be

considerably less than for message switching, despite the inclusion of a header in each packet rather than in each message.

4.5 WAN Virtual Circuits

A virtual circuit is a logical circuit created within a shared network between two network devices. Two types of virtual circuits exist: switched virtual circuits (SVCs) and permanent virtual circuits (PVCs).

SVCs are virtual circuits that are dynamically established on demand and terminated when transmission is complete. Communication over an SVC consists of three phases: circuit establishment, data transfer, and circuit termination. The establishment phase involves creating the virtual circuit between the source and destination devices. Data transfer involves transmitting data between the devices over the virtual circuit, and the circuit termination phase involves tearing down the virtual circuit between the source and destination devices. SVCs are used in situations in which data transmission between devices is irregular, largely because SVCs increase bandwidth used due to the circuit establishment and termination phases, but they decrease the cost associated with constant virtual circuit availability.

PVC is a permanently established virtual circuit that consists of one mode: data transfer. PVCs are used in situations in which data transfer between devices is constant. PVCs decrease the bandwidth use associated with the establishment and termination of virtual circuits, but they increase costs due to constant virtual circuit availability. PVCs are generally configured by the service provider when an order is placed for service.

4.6 WAN Dialup Services

Dialup services offer cost-effective methods for connectivity across WANs. Two popular dialup implementations are dial-on-demand routing (DDR) and dial backup.

DDR is a technique whereby a router can dynamically initiate a call on a switched circuit when it needs to send data. In a DDR setup, the router is configured to

initiate the call when certain criteria are met, such as a particular type of network traffic needing to be transmitted. When the connection is made, traffic passes over the line. The router configuration specifies an idle timer that tells the router to drop the connection when the circuit has remained idle for a certain period.

Dial backup is another way of configuring DDR. However, in dial backup, the switched circuit is used to provide backup service for another type of circuit, such as point-to-point or packet switching. The router is configured so that when a failure is detected on the primary circuit, the dial backup line is initiated. The dial backup line then supports the WAN connection until the primary circuit is restored. When this occurs, the dial backup connection is terminated.

4.7 WAN Devices

WANs use numerous types of devices that are specific to WAN environments. WAN switches, access servers, modems, CSU/DSUs, and ISDN terminal adapters are discussed in the following sections. Other devices found in WAN environments that are used in WAN implementations include routers, ATM switches, and multiplexers.

4.7.1 WAN Switch

A WAN switch is a multiport internetworking device used in carrier networks. These devices typically switch such traffic as Frame Relay, X.25, and SMDS, and operate at the data link layer of the OSI reference model. Figure 4.8 illustrates two routers at remote ends of a WAN that are connected by WAN switches.

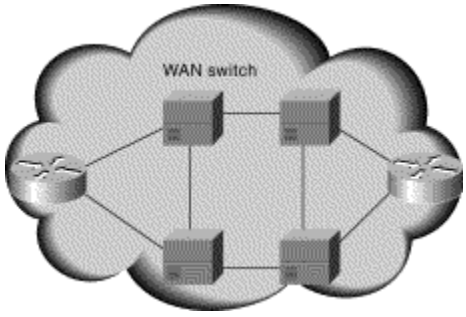


Figure 4.8 WAN Switches

4.7.2 Access Server

An access server acts as a concentration point for dial-in and dial-out connections. Figure 4.9 illustrates an access server concentrating dial-out connections into a WAN.

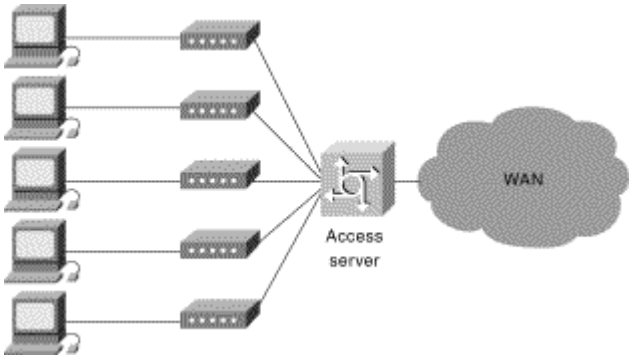


Figure 4.9 Role of Access Server

4.7.3 Modem

A modem is a device that interprets digital and analog signals, enabling data to be transmitted over voice-grade telephone lines. At the source, digital signals are converted to a form suitable for transmission over analog communication facilities. At the destination, these analog signals are returned to their digital form. Figure 4.10 illustrates a simple modem-to-modem connection through a WAN.

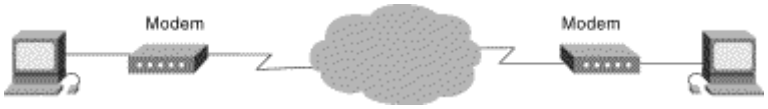


Figure 4.10 Modem to Modem Connection

4.7.4 CSU/DSU

A channel service unit/digital service unit (CSU/DSU) is a digital-interface device used to connect a router to a digital circuit like a T1 or DS1. Digital signal 1 (DS1, also known as T1, sometimes "DS-1") is a T-carrier signaling scheme devised by Bell Labs. DS1 is a widely used standard in telecommunications in North America and Japan to transmit voice and data between devices. E1 is used in place of T1 outside of North America and Japan. Technically, DS1 is the transmission protocol used over a physical T1 line; however, the terms "DS1" and "T1" are often used interchangeably. The CSU/DSU also provides signal timing for communication between these devices. Figure 4.11 illustrates the placement of the CSU/DSU in a WAN implementation.

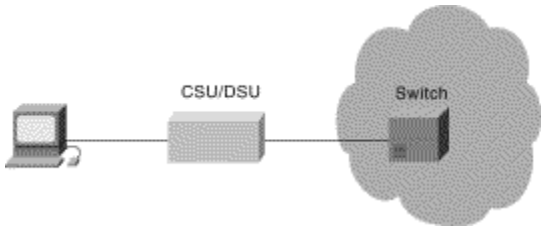


Figure 4.11 CSU/DSU Stands between the switch and terminal

4.7.5 ISDN Terminal Adapter

An ISDN terminal adapter is a device used to connect ISDN Basic Rate Interface (BRI) connections to other interfaces, such as EIA/TIA-232 on a router. A terminal adapter is essentially an ISDN modem, although it is called a terminal adapter because it does not actually convert analog to digital signals. Figure 4.12 illustrates the placement of the terminal adapter in an ISDN environment.

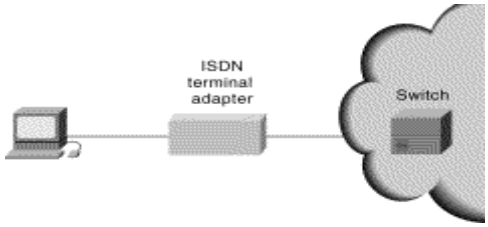


Figure 4.12 Terminal Adapter in an ISDN Environment

4.8 LAN Switch

A typical network consists of:

- nodes, or computers
- a medium for connection, either wired or wireless
- special network equipment, such as routers or hubs

In the case of the Internet, these pieces work together to allow your computer to send information to another computer. The other computer can be on the other side of the world.

Switches are a fundamental part of most networks. Switches enable several users to send information over a network. Users can send the information at the same time and do not slow each other down. Just like routers allow different networks to communicate with each other, switches allow different nodes of a network to communicate directly with each other. A node is a network connection point, typically a computer. Switches allow the nodes to communicate in a smooth and efficient manner.

There are many different types of switches and networks. Switches that provide a separate connection for each node in a company internal network have the name LAN switches. Essentially, a LAN switch creates a series of instant networks that contain only the two devices that communicate with each other at that particular moment.

Hubs provide an easy way to scale up and shorten the distance that the packets must travel to get from one node to another. But hubs do not break up the actual network into discrete segments. Switches handle this job.

Think of a hub as a four-way intersection where all vehicles have to stop. If more than one car reaches the intersection at one time, the cars must wait for a turn to proceed. But a switch is like a cloverleaf intersection. Each car can take an exit ramp to get to the destination without the need to stop and wait for other traffic to pass. Now imagine this scenario with a dozen or even a hundred roads that intersect at a single point. The wait and the potential for a collision increase significantly if every car has to check all the other roads before the car proceeds.

Imagine that you can take an exit ramp from any one of those roads to the road of your choice. This ability is what a switch provides for network traffic.

There is a vital difference between a hub and a switch; all the nodes that connect to a hub share the bandwidth, but a device that connects to a switch port has the full bandwidth alone. For example, consider 10 nodes that communicate with use of a hub on a 10 Mbps network. Each node can only get a portion of the 10 Mbps if other nodes on the hub want to communicate as well. But, with a switch, each node can possibly communicate at the full 10 Mbps. Consider the road analogy. If all the traffic comes to a common intersection, the traffic must share that intersection. But a cloverleaf allows all the traffic to continue at full speed from one road to the next.

In a fully switched network, switches replace all the hubs of an Ethernet network with a dedicated segment for every node. These segments connect to a switch, which supports multiple dedicated segments. Sometimes the number of segments reaches the hundreds. Since the only devices on each segment are the switch and the node, the switch picks up every transmission before the transmission reaches another node. The switch then forwards the frame over the appropriate segment. Since any segment contains only a single node, the frame only reaches the intended recipient. This arrangement allows many conversations to occur simultaneously on a network that uses a switch.

Switching allows a network to maintain full-duplex Ethernet. Before switching existed, Ethernet was half duplex. Half duplex means that only one device on the network can transmit at any given time. In a fully switched network, nodes only communicate with the switch and never directly with each other. In the road analogy, half duplex is similar to the problem of a single lane, when road construction closes one lane of a two-lane road. Traffic attempts to use the same lane in both directions. Traffic that comes one way must wait until traffic from the other direction stops in order to avoid collision.

Fully switched networks employ either twisted pair or fiber-optic cable setups. Both twisted pair and fiber-optic cable systems use separate conductors to send and receive data. In this type of environment, Ethernet nodes can forgo the

collision detection process and transmit at will; these nodes are the only devices with the potential to access the medium. In other words, the network dedicates a separate lane to traffic that flows in each direction. This dedication allows nodes to transmit to the switch at the same time that the switch transmits to the nodes. Thus, the environment is collision-free. Transmission in both directions also can effectively double the apparent speed of the network when two nodes exchange information. For example, if the speed of the network is 10 Mbps, each node can transmit at 10 Mbps at the same time.

Most networks are not fully switched because replacement of all the hubs with switches is costly. Instead, a combination of switches and hubs create an efficient yet cost-effective network. For example, a company can have hubs that connect the computers in each department and a switch that connects all the department-level hubs together.

A switch has the potential to radically change the way that the nodes can communicate with each other. But what makes a switch different than a router? Switches usually work at Layer 2 (Data or Datalink) of the Open System Interconnection (OSI) reference model with use of MAC addresses. Routers work at Layer 3 (Network) with Layer 3 addresses. The routers use IP, Internetwork Packet Exchange (IPX), or Appletalk, which depends on the Layer 3 protocols that are in use. The algorithm that switches use to decide how to forward packets is different than the algorithms that routers use to forward packets. One difference in the algorithms is how the device handles broadcasts. On any network, the concept of a broadcast packet is vital to the operability of the network. Whenever a device needs to send out information but does not know to whom to send the information, the device sends out a broadcast. For example, every time a new computer or other device comes onto the network, the device sends out a broadcast packet to announce the entry. The other nodes, such as a domain server, can add the device to the browser list. The browser list is like an address directory. Then, the other nodes can communicate directly with that device. A device can use broadcasts to make an announcement to the rest of the network at any time.

A hub or a switch passes along any broadcast packets that the device receives to all the other segments in the broadcast domain. But a router does not pass along broadcast packets. Think about the four-way intersection again. In the analogy, all the traffic passes through the intersection, despite the direction of travel. Now, imagine that this intersection is at an international border. In order to pass through the intersection, you must provide the border guard with the specific address to which you are going. If you do not have a specific destination, the guard does not let you pass. A router works in a similar way. If a data packet does not have the specific address of another device, the router does not let the data packet pass. This restriction keeps networks separate from each other, which is good. But, when you want to talk between different parts of the same network, the restriction is not good. Switches can overcome this restriction.

LAN switches rely on packet switching. The switch establishes a connection between two segments and keeps the connection just long enough to send the current packet. Incoming packets, which are part of an Ethernet frame, save to a temporary memory area. The temporary memory area is a buffer. The switch reads the MAC address that is in the frame header and compares the address to a list of addresses in the switch lookup table. In a LAN with an Ethernet basis, an Ethernet frame contains a normal packet as the payload of the frame. The frame has a special header that includes the MAC address information for the source and destination of the packet.

Switches use one of three methods for routing traffic:

- Cut-through
- Store and forward
- Fragment-free

Cut-through switches read the MAC address as soon as a packet is detected by the switch. After storing the six bytes that make up the address information, the switches immediately begin to send the packet to the destination node, even though the rest of the packet is coming into the switch.

A switch that uses store and forward saves the entire packet to the buffer and checks the packet for Cyclic Redundancy Check (CRC) errors or other problems.

If the packet has an error, the packet is discarded. Otherwise, the switch looks up the MAC address and sends the packet on to the destination node. Many switches combine the two methods by using cut-through until a certain error level is reached, then changing over to store and forward. Very few switches are strictly cut-through because this provides no error correction.

A less common method is fragment-free. Fragment-free works like cut-through, but stores the first 64 bytes of the packet before sending the packet on. The reason for this is that most errors and all collisions occur during the initial 64 bytes of a packet.

LAN switches vary in physical design. Currently, there are three popular configurations in use:

Shared-memory—The switch stores all incoming packets in a common memory buffer that all the switch ports (input/output connections) share. Then, the switch sends the packets out the correct port for the destination node.

Matrix—This type of switch has an internal grid with which the input ports and the output ports cross each other. When the switch detects a packet on an input port, the switch compares the MAC address to the lookup table to find the appropriate output port. The switch then makes a connection on the grid where these two ports intersect.

Bus-architecture—Instead of a grid, an internal transmission path (common bus) is shared by all the ports using time division multiplex access (TDMA). A switch with this configuration dedicates a memory buffer to each port. There is an application-specific integrated circuit (ASIC) to control the internal bus access.

4.9 Bridge

It interconnects LAN segments at the Network Interface layer level and forwards frames between them. A bridge performs the function of a MAC relay, and is independent of any higher layer protocol (including the Logical Link protocol). It provides MAC layer protocol conversion, if required. Examples of bridges are:

- A PS/2 running the IBM Token-Ring Network Bridge program
- The IBM 8229 LAN bridge

A bridge can be said to be transparent to IP. That is, when a host sends an IP datagram to another host on a network connected by a bridge, it sends the datagram directly to the host and the datagram "crosses" the bridge without the sending host being aware of it. The features of Bridge are-

- Hardware device
- Connects two LAN segments
- Forwards frames
- Does not forward noise or collisions
- Learns addresses and filters
- Allows independent transmission

4.9.1 Operation of a Bridge

The simplest type of bridge, and that most frequently used is the Transparent Bridge (meaning that the nodes using a bridge are unaware of its presence). The bridge therefore has to forward (receive and subsequently transmit) frames from one LAN (e.g. LAN A below) to another (e.g. LAN B). Obviously, the bridge could forward all frames, but then it would behave rather like a repeater; it would be much smarter if the bridge only forwarded frames which need to travel from one LAN to another. To do this, the bridge needs to learn which computers are connected to which LANs. More formally, it needs to learn whether to forward to each address.

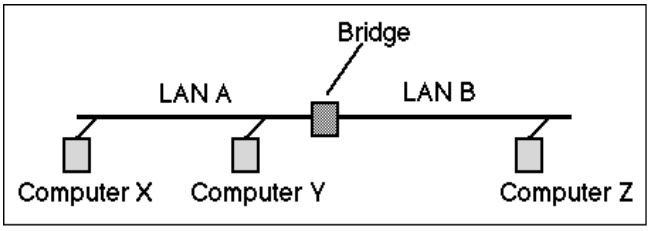


Figure 4.13 A bridge connecting two LAN segments (A and B)

To learn which addresses are in use, and which ports (interfaces on the bridge) theory are closest to, the bridge observes the headers of received Ethernet frames. By examining the MAC source address of each received frame, and recording the port on which it was received, the bridge may learn which

addresses belong to the computers connected via each port. This is called "learning". In the figure above, consider three computers X,Y,Z. Assume each sends frames to the other computers. The source addresses X,Y are observed to be on network A, while the address of computer Z will be observed to be on network B.

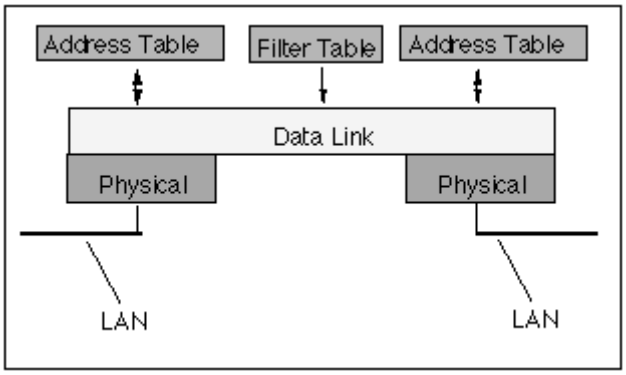


Figure 4.14 A bridge stores the hardware addresses observed from frames received by each interface and uses this information to learn which frames need to be forwarded by the bridge

The learned addresses are stored in the an interface address table associated with each port (interface). Once this table has been setup, the bridge examines the destination address of all received frames, it then scans the interface tables to see if a frame has been received from the same address (i.e. a packet with a source address matching the current destination address). Three possibilities exist:

- If the address is not found, no frames have been received from the source. The source may not exist, or it may not have sent any frames using this address. (The address may also have been deleted by the bridge because the bridge software was recently restarted, ran short of address entries in the interface table, or deleted the address because it was too old). Since the bridge does not know which port to use to forward the frame, it will send it to all output ports, except that on which it was received. (It is clearly unnecessary to send it back to the same cable segment from which it was received, since any other computer/bridges on this cable must already have received the packet.) This is called flooding.

- If the address is found in an interface table and the address is associated with the port on which it was received, the frame is discarded. (It must already have been received by the destination.)
- If the address is found in an interface table and the address is not associated with the port on which it was received, the bridge forwards the frame to the port associated with the address.

Packets with a source of X and destination of Y are received and discarded, since the computer Y is directly connected to the LAN A, whereas packets from X with a destination of Z are forwarded to network B by the bridge.

4.9.2 Broadcast and Multicast

Bridges forward a broadcast frame out of all connected ports except that on which the frame was received. The normal action for multicast frame is to treat them as broadcast frame. This is clearly suboptimal, since a bridge may send multicast frames to parts of the network for which there are no interested receivers. Some bridges implement extra processing to control the flooding of multicast frames.

4.9.3 Managing the Interface Tables

A bridge may implement an interface table using a software data structure or use a Contents Addressable Memory (CAM) chip. In either case, the size of the table is finite, and usually constrained to 1000's - 10 000's of entries. In a large LAN this may be a limit. To help keep the table small, most bridges maintain a check of how recently each address was used. Addresses which have not been used for a long period of time (e.g. minutes) are deleted. This has the effect of removing unused entries, but if the address is again used, before a frame is received from the same source, it will require the frame to be flooded to all ports.

A useful side effect of deleting old addresses is that the bridge interface table records only working MAC addresses. If a NIC stops sending, its address will be deleted from the table. If the NIC is subsequently reconnected, the entry will be restored, but if the connection is made to another port (the cable is changed) a

different (updated) entry will be inserted corresponding to the actual port associated with the address. (The bridge always updates the interface table for each source address in a received MAC frame, therefore even if a computer changes the point at which it is connected without first having the interface table entry removed, the bridge will still update the table entry).

4.9.4 Filter Tables

In some bridges, a system administrator may override the normal forwarding by inserting entries in a filter table to inhibit forwarding between different work groups (for example to provide security for a particular set of MAC addresses). The filter table contains a list of source or destination addresses. Frames which match entries in the filter table will only be forwarded to specific configured ports.

4.10 Router

Router interconnects networks at the inter-network layer level and routes packets between them. The router must understand the addressing structure associated with the networking protocols it supports and take decisions on whether, or how, to forward packets. Routers are able to select the best transmission paths and optimal packet sizes. The basic routing function is implemented in the IP layer of the TCP/IP protocol stack. Therefore any host or workstation running TCP/IP may be used as a router. However, dedicated routers such as the IBM 6611 Network Processor provide much more sophisticated routing than the minimum function implemented by IP. Because IP provides this basic routing function, the term “IP router”, is often used. Other, older, terms for router are “IP gateway”, “Internet gateway” and “gateway”. The term gateway is now normally used for connections at a higher level than the router level.

A router can be said to be visible to IP. That is, when a host sends an IP datagram to another host on a network connected by a router, it sends the datagram to the router and not directly to the target host. In addition to identifying networks and providing connectivity, routers also provide other functions:

- Routers do not forward Layer 2 broadcast or multicast frames.

- Routers attempt to determine the optimal path through a routed network based on routing algorithms.
- Routers strip Layer 2 frames and forward packets based on Layer 3 destination addresses.
- Routers map a single Layer 3 logical address to a single network device; therefore, routers can limit or secure network traffic based on identifiable attributes within each packet. These options, controlled via access lists, can be applied to inbound or outbound packets.
- Routers can be configured to perform both bridging and routing functions.
- Routers provide connectivity between different virtual LANs (VLANs) in a switched environment.
- Routers can be used to deploy quality of service parameters for specified types of network traffic.

4.11 Gateway

Gateway interconnects networks at higher levels than bridges or routers. A gateway usually supports address mapping from one network to another, and may also provide transformation of the data between the environments to support end-to-end application connectivity. Gateways typically limit the interconnectivity of two networks to a subset of the application protocols supported on either one. For example, a VM host running TCP/IP may be used as an SMTP (Simple Mail Transfer Protocol) mail gateway.

VM (often: VM/CMS) refers to a family of IBM virtual machine operating systems used on IBM System/370, System/390, zSeries, and System z9 IBM mainframes and compatible systems.

Note: The term “gateway”, when used in this sense, is not synonymous with “IP gateway”.

A gateway can be said to be opaque to IP. That is, a host cannot send an IP datagram through a gateway: it can only send it to a gateway. The higher-level protocol information carried by the datagrams is then passed on by the gateway using whatever networking architecture is used on the other side of the gateway.

Closely related to routers and gateways is the concept of a firewall or firewall gateway which is used to restrict access from the Internet to a network or a group of networks controlled by an organization for security reasons.

4.12 Repeaters

Repeaters operate within the physical layer of the OSI reference model and regenerate the signal. Repeaters are used in LANs, MAN and WANs. They may be used to provide more flexibility in design of a network or to extend the distance over which a signal may travel down a cable. One example of a repeater is an Ethernet Hub.

4.13 Summary

Circuit Switching allows the entire circuit for a connection while packet switching allows the multiple packets of different connections to pass on a single network. Switches break up the actual network into discrete segments. Bridges interconnects LAN segments and forwards frames between them. The router uses the information held in the network layer header (i.e. IP header) to decide whether to forward each received packet, and which network interface to use to send the packet.

4.14 Self Assessment Questions

Q1. Explain the differences between circuit switching network and packet switching network.

Q2. Explain different WAN devices.

Q3. What are the advantages of using switches over hubs?

Q4. Explain the operation of bridges.

Q5. Explain the layers on which following devices work-

i) Hub

ii) Switch

iii) Bridge

iv) Router

4.15 References

1. Computer Networks, Andrew S. Tanenbaum, Prentice Hall, 2002
2. Computer Networks, Randall Rustin, Pearson Education, 2000.
3. Computer Networks, V.S.Bagad, L.A.Dhotre, Technical Publications

Paper Code: MCA- 405	**Author: Parvinder Singh**
Paper Name: Computer Networks-II	**Vetter: Dr. Sudhir Batra**
Lesson Number: 05	

Internet Security

Structure

5.0 Objective

The objective of this chapter is to study various security concerns and dangers on internet. The measures implemented on internet to make it a secure place, are depicted. This chapter also describes the Firewall designs, IPv6 and internet tools available.

5.1 Introduction

Internet Security is one of the major issues with the Internet because it is public domain. The public nature of the Internet can cause security concerns that don't exist for private intranet or dial-up applications. Because packets pass through machines over which you have no control, someone can potentially see confidential information. Any hacker with a network data scope can get credit card numbers, Social Security numbers, and other confidential information from your transmissions. We need to design for these potential security leaks.

5.2 Passing through Multiple Machines

Your transactions have the potential to pass through many computers and other devices on their way between the client and the host. On most UNIX systems, you can issue the traceroute command to see this routing. Most of these machines are acting only as routers, but they are points where your signal can be intercepted and decoded. Anyone with a scope on any of the devices through which your information passes can trap that information. Things like Social Security numbers (999-99-9999) and credit card numbers have patterns that can be detected by automated search programs. An unscrupulous person can place one of these programs on a device routing packets along the Internet, let it work for a period of time, and then take a leisurely look at the data that it traps.

5.2.1 E-Mail Example

E-mail can be even more vulnerable to this type of piracy, because mail travels as plain text in a format that's easy to read, and full messages are stored and forwarded by post office machines. Although most of us don't like to look at them,

and many mail readers filter them, mail headers can tell you a lot about the machines on which your mail rests. Take a look at a message header:

Received: from ns2.eds.com by mail5.netcom.com (8.6.12/Netcom)

id NAA01582; Wed, 24 Jan 1996 13:21:17 -0800

Received: by ns2.eds.com (hello)

id QAA07685; Wed, 24 Jan 1996 16:21:40 -0500

Received: by nnsp.eds.com (hello)

id QAA26247; Wed, 24 Jan 1996 16:19:58 -0500

Received: from target2.sssc.slg.eds.com by dsscsun1.dssc.slg.eds.com

(5.0/SMI-SVR4)

id AA00143; Wed, 24 Jan 1996 15:18:57 -0600

Received: from rfbpc (rfbpc.sssc.slg.eds.com [198.132.57.4])

by target2.sssc.slg.eds.com

The details of this heading information aren't important for this discussion. The important thing is the fact that this piece of mail rested on four machines are not under control. At each of these points, message is simply part of a larger text file. Anyone with the proper security clearance (or anyone who can hack into that machine and obtain that clearance) can read the message. The headings are read from the bottom to the top:

- The mail originated on PC (rfbpc).
- The mail was passed to the post office machine on LAN (target2).
- The mail was forwarded by post office to division's mail handler (dsscsun1).
- The division mail post office passed the mail to corporate firewall (nnsp).
- The mail passed to the corporate post office outside the firewall (ns2).
- Finally, the mail was delivered to the post office on the Internet service provider (mail5).

Incidentally, the mail passed through several machines that aren't listed in this heading. Remember that traceroute? Mail packets have to pass through several machines on which they don't rest, making them vulnerable to snooping.

What does this mean to your application? If you're passing sensitive, private, or confidential information, consider encryption for your application.

5.3 Encryption

Encryption is a technique for scrambling and unscrambling information. The unscrambled information is called clear-text and the scrambled information is called cipher-text. Once data is encrypted, it can be sent via e-mail messages or stored just as any other data. To read the data, the recipient must decrypt, or decipher, it into a readable form. To encrypt the data the originator of the data applies an encryption key, secret values that computers use along with complex mathematical formulas to encrypt messages. The recipient of the data then uses an encryption key to decrypt the data.

There are two basic types of encryption, private key and public key. With private key encryption, symmetric encryption, both the originator and recipient use the same encryption key to encrypt and decrypt the data. Public key encryption uses two encryption keys: a public key known to everyone and a private key known to the receiver only. To decode an encrypted message, a computer must use the public key, provided by the originating computer and its own private key.

Many browsers include encryption software that allows the user to encrypt e-mail messages or other documents. Secure Socket Layer (SSL) is one of the more popular Internet encryption methods, which provides two-way encryption along the entire route data travels to and from a computer. Web pages those use SSL begin with the https protocol, instead of http protocol. To check if the communication is carried over a secure channel that uses SSL, the user should look for a padlock icon on the browser and make sure that the URL is in the form of https:// as opposed to http://. Before entering any sensitive data, such as a credit card number or Social Security number, it is important to verify that the user's computer is communicating with the right server and not an imposter that is trying to steal personal information. To verify this, the user should check the authentication certificate, which could be accessed by double-clicking on the padlock icon.

Many types of encryption can be used to protect your transactions. Several Web browsers and hosts are “secure” in that they encrypt information passing between them. The extent to which we want to use encryption in your application will depend on the sensitivity of the information and the cost of encryption.

Of course, if you are writing your own application in which you will provide both the client and server modules, you can provide your own custom encryption schemes. One caution about using encryption such as that used by products like Pretty Good Privacy. These schemes are controlled by the U.S. Federal Government, which has some restrictions against exporting encryption technology overseas. Be sure to check out this issue before committing your application to specific technology or standards.

5.3.1 Secure Web Servers

If you are designing an application that will be hosted by a Web server, consider placing the application on a secure Web server. These servers establish a secure connection with the client browser and encrypt all information that passes between them. The Netscape Commerce Server, for example, uses Secure Sockets Layer (SSL) to encrypt pages during transmission.

5.3.2 Encrypting Sensitive Information

Even if you choose not to encrypt entire transmissions, never send an unencrypted password, Social Security number, credit card number, or other sensitive information over the Internet. This data can be encrypted easily by the host CGI interface program, even if you implement your program using a commercial Web hosting program. Implementing encryption at the client end of the application is more difficult if you don't rely on the encryption capabilities of the commercial server/client. Java or some other plug-in application needs to be used to encrypt the sensitive information prior to transmission.

5.3.3 Encrypting or Password-Protecting Documents

If you are going to transmit documents over the Internet, such as word processing documents, you can use the capabilities of the applications that create the documents to encrypt or password-protect the documents. For example, both Microsoft Word for Windows and Microsoft Excel can provide file-sharing passwords that must be entered before a document can be accessed.

You might also want to use the capacity of compression programs such as PKZIP to password-protect files they have compressed. With this system, even if some hacker manages to intercept a file, she will have to work hard to read it.

Following are some thoughts about using passwords:

- Use the longest password you can to protect your documents.
- Don't use common words or phrases. They are easier to remember, but also easier to crack. Random combinations of letters and numbers are best.
- Change passwords periodically. That is, if you send several documents over a period of time, make sure that you change the password at least every 30-60 days. This strategy lessens the chance that the password will be compromised.
- Don't send passwords over the Internet; transmit them via a secure means.

5.3.4 Unsecured Request, Secure Response

In the case of especially sensitive information, you can allow requests to come to your application via the public Internet. However, you might want to return the requested information via a secure medium. For example, you could allow customers to request information via the Internet and then use fax-back facilities to fax the information to their machines.

5.3.5 Verifying the Correct Client

Another difficulty in dealing with connectionless protocols is that you might need to verify that the client you are talking to is the one you think it is. Luckily, some techniques are available, as described in the following sections.

5.3.6 Trusted Addresses

Your application might accept socket connections only from “trusted” TCP/IP addresses. Web browsers send the name of the machine in the SERVER_NAME field and the address of the remote in the REMOTE_ADDRESS field. Be aware that these fields can be faked, but they can be used in combination with user IDs and passwords to provide additional security.

5.3.7 User IDs and Passwords

Your application might ask the client for a user ID and password. For applications with custom clients, the user ID can be programmed into the client before distribution and the user can be required to enter a specific password to verify her identity. In addition, you can limit user IDs to specific TCP/IP addresses and refuse to serve ID/address pairs those don't match.

5.4 Cookies

If your application uses commercial browsers, you can take advantage of the capacity of some browsers-for example, Netscape or Microsoft's Internet

Explorer-to store information on the client machine; that information can be returned to the server when a specific host path is requested.

CGI scripts can set data at the client's browser; this information is called a magic cookie. When a browser makes a request for a page, it sends its cookie (if it has one set) to the server along with the request. If this is the first time that this particular machine has been used to access your application, it will need to set default configurations or provide a form on which the customer can provide required information.

A magic cookie is made up of several parts:

- URLs for which the cookie will be sent
- PATH for which the cookie will be sent in the above host domain
- DATE when the browser will delete its cookie
- SECURE flag that tells the browser to send its cookie only if it has a secure connection to the server

Using this capability, you could transmit a user ID to the client and then retrieve it on subsequent visits by this client. You can match the returned cookie to security information entered by the human being on-screen as an additional security precaution.

Cookies can also be a convenient way to customize your application for a particular client; for example, when you are transmitting a page in a foreign language for international clients. Once a customer has visited your site, you can recognize the customer from his cookie, and automatically customize the page returned to him.

5.5 Firewall Concepts

The concept that stands behind the firewall approach is to allow local users to enjoy full network services within their local network and some useful services provided by the Internet while controlling outsiders' access to the local network resources. Firewall approach achieves security by isolating a specific segment of Internet topology(further Local Network) from the rest of the Internet and controlling all the traffic that comes to and leaves the Local Network.

To control the network traffic each connection of Local Network to the Internet is equipped with a firewall. Firewall's goal is to inspect and control all the traffic between the Local Network and the Internet. The traffic must be handled in such a way that all potentially “dangerous” traffic be detected and dropped and if necessary logged. What traffic is “dangerous” for the Local Network is determined by the Security Policy adopted for the site. Figure 5.1 shows Local Network without Firewall.

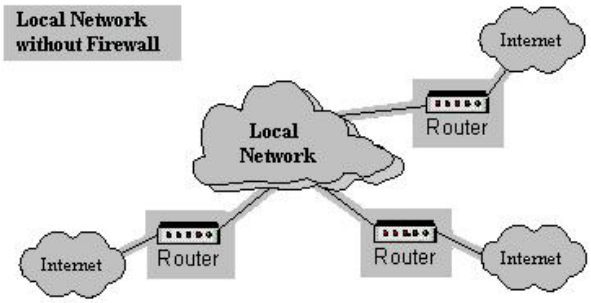


Figure 5.1(Local Network without Firewall)

5.5.1 Why Firewalls?

The result of firewalling the Local Network can be viewed as follows. In the case of Local Network directly connected to the Internet without any firewall, the entire network is subject to a attack. Consider a large organization with thousands of hosts. If every host is allowed to communicate directly with the outside world, the attackers will find the weakest of the hosts and penetrate it. If one of the hosts is penetrated it is not difficult to penetrate all the other hosts on the network using the resources of that compromised host. Practical experience shows that it is very difficult to ensure that every host on the network is secure. One badly chosen password and all the network security can be compromised. On the other hand if Local Network is guarded by the firewall there is direct access only to selected subset of hosts and the zone of risk is often reduced to the firewall itself or a selected subset of hosts on the network. In some sense firewall is not so

much a security solution as they are a response to the engineering/administration problem: configuring a large number of hosts systems for good security.

As was mentioned above firewall must inspect all the packets that come to and leave the Local Network and filter out those packets that do not conform to the security policy adopted for the Local Network.

Remember the ISO seven layers protocol model. The packet inspection can take place on any of the layers. But packet inspection is most commonly implemented at Application layer by Application layer firewalls and at Network layer by Network layer firewalls.

Communication Layers
Application
Presentation
Session
Transport
Network
Data Link
Physical

Figure 5.2 (ISO Model)

When talking about TCP/IP protocol suite the Application layer firewalls are commonly called Application Gateways or Proxies(further Proxies) and Network layer firewalls Filtering Routers or Screening Routers(further Filtering Routers).

5.5.2 How packets are filtered out?

Remember the function of ordinary IP router. It receives IP datagram extracts destination IP address and consults the routing table for next hop for this datagram.

As its name indicates Filtering Router in addition to routing function performs a filtering of the packets it receives, that is before consulting the routing table it must decide whether this packet should be forwarded towards its destination.
The filtering decision is made according to the Access Control List (further ACL) associated with physical interface the packet came through.
ACL consists of the entries. Each entry specifies values for particular header fields and action to be taken if arrived packet matches these values.
Each arrived packet is matched successively against the entries in the ACL if the match occurs the action is taken. Figure 5.3 shows the steps of filtering decisions.

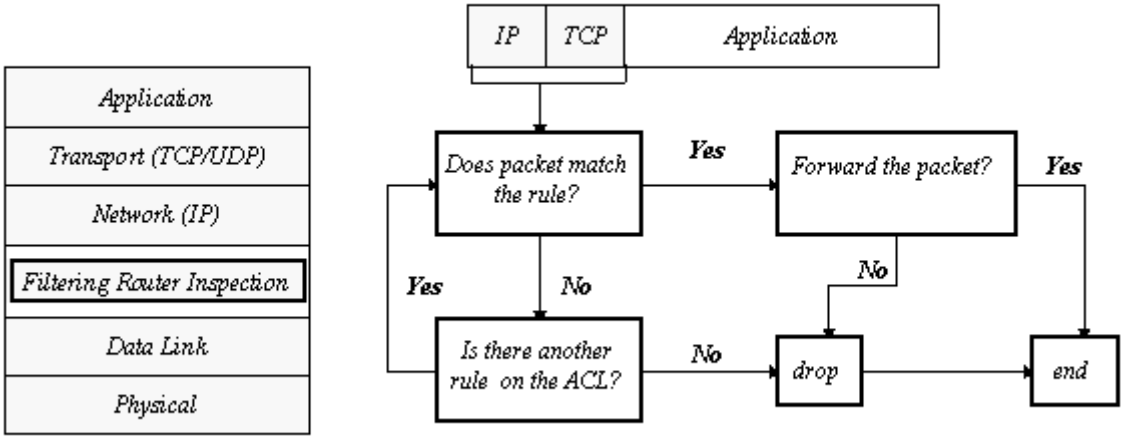


Figure 5.3 (Steps of Filtering Decision)

The question arises: “What is done with the packet that does not match any entry in the ACL?”
In this situation two different approaches may be adopted:

- That which is not expressively permitted is prohibited, that is these packets will be dropped by the Filtering Router;
- That which is not expressively prohibited is permitted, that is these packets will forwarded by the Filtering Router;

5.5.3 What information is used for filtering decision?

The portions that are parsed by the filtering router are IP header, and transport protocol header whether TCP or UDP. Therefore the header fields that can be used in ACL entries are:

- Source IP address (IP header),
- Destination IP address (IP header),
- Protocol Type (IP header, specifies whether the data encapsulated in the IP datagram belongs to TCP, UDP or ICMP protocol),
- Source port (TCP or UDP header),
- Destination port (TCP or UDP header),
- ACK. bit (TCP header, this bit specifies whether the packet is the acknowledgment for received TCP packet);

5.5.4 Firewall Components: Application Level Gateways

5.5.4.1 Proxy

Proxy is -The agency, function, or power of a person authorized to act as the deputy or substitute for another;

The idea that stands behind Proxy component of a firewall design is not to allow direct TCP (UDP) connection between client software on the Local Network and server software on the Internet or vice versa (the client software on the Internet and server software on the Local Network). Instead the direct connection is broken into two separate connections. The proxy program is acting as an intermediate. Operation of application level gateway is shown in figure 5.4.

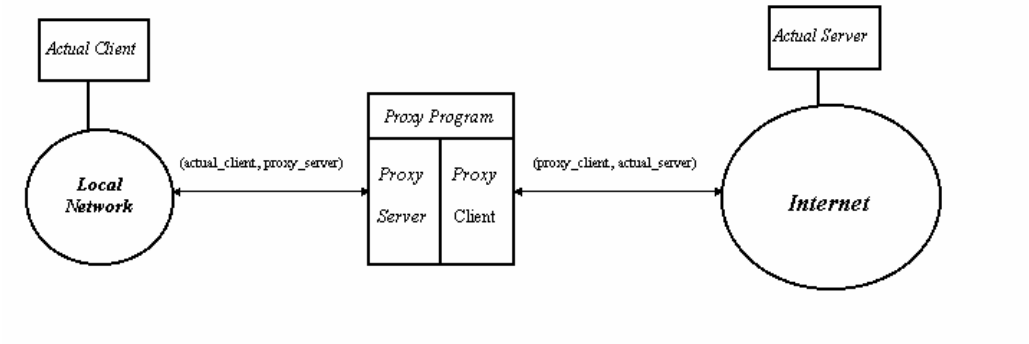


Figure 5.4 (Operation of Application Level Gateway)

As it relays the traffic between actual client and actual server it does checks and access controls that typical client and server software do not support.

5.5.4.2 How proxy program works?

Proxy program must implement enough of the client and server part of application protocol to accomplish the following:

- Accept client sessions and appear to them as a server;
- Receive from the client software the name of the actual server;
- Contact the actual server and appear to it as a client;
- Relay all the data from the client to a server;
- Perform access control function, that is according to Security Policy chosen for a site it must reject potentially dangerous connections.

5.5.4.3 Enforcement of external client to proxy server

First of all the external hosts should not know about the Local Network topology and thus should not know the IP address or name of the host machine on the Local Network on which a specific server may be located. External hosts should only know about the Application Gateway machines.

But if external hosts does have this information it could try to contact the internal server. To prevent this the IP connectivity between Local Network and the

Internet must be broken. This can be achieved in several ways please refer Firewall Architectures part of this document.

5.5.4.4 Enforcement of internal client to proxy server

Once more the IP connectivity between Local Network and the Internet must be broken.

In addition to this the client software on the Local Network must know how to contact the proxy server instead of the actual server on the Internet. To accomplish this two proxy technologies exist: classical proxy technology and transparent proxy technology.

In classical proxy technology either the client software is modified or the user is instructed to follow special setup procedures in order to make call to the actual server through the proxy server.

In transparent proxy technology the routing tables of the Local Network are configured in such a way that all the packets destined for the external servers come to the Application Gateway machine and proxy program knows how to intercept these packets and to form two connections (actual_client, proxy_server) and (proxy_client, actual_server).

5.5.5 Firewall Architectures

5.5.5.1 Screening Router Architecture

In this architecture a firewall consists of nothing more than a screening router. Figure 5.5 shows this architecture. Host on the Local Network and hosts on the Internet are allowed to communicate directly. The communication is restricted to the type that is allowed by a screening router. The security of the whole Local Network depends on the correct ACL of the router and on the amount of services permitted.

5.5.5.2 Dual-Homed Host Architecture

In this architecture a firewall consists of Dual-Homed Host machine (machine having two or more IP addresses each for specific physical port). One port of the

machine connects to the Local Network and the other port/ports connects to the Internet. The IP datagram forwarding is turned off on the Dual-Homed Host machine, thus there is no direct TCP/IP connection between the Local Network and the Internet. Figure 5.6 shows this architecture.

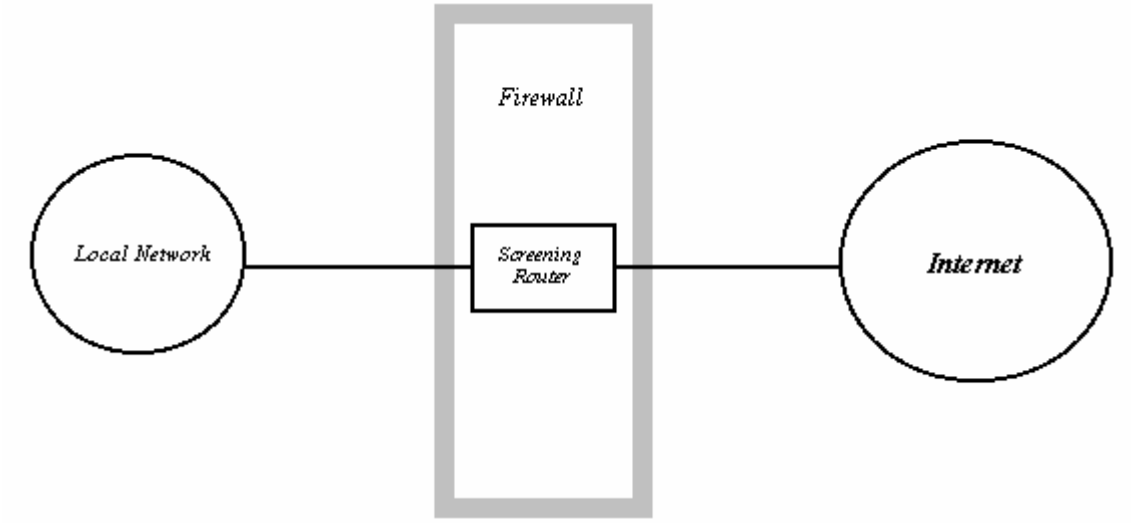


Figure 5.5

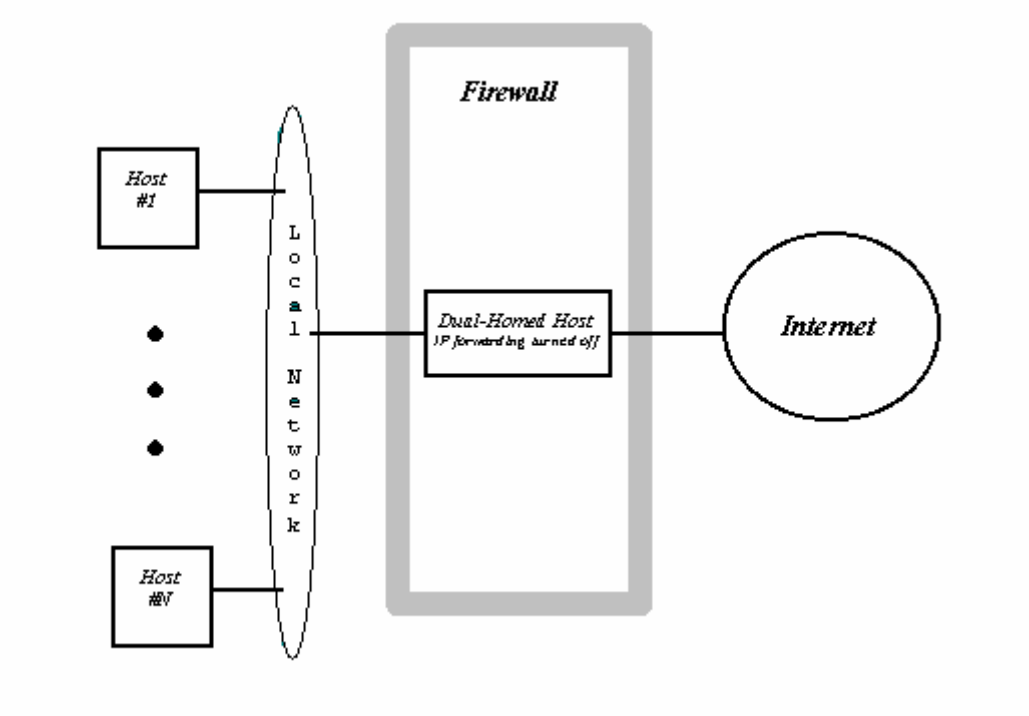


Figure 5.6

You permit communication between Local Network and the Internet in either of two ways:

- Users on the Local Network are given accounts on the Dual-Homed Host machine. In order to use Internet services the must rlogin on the Dual-Homed Host machine. The fact that you allow accounts on the machine weakens its security greatly (it now depends on each user and user that have access to it, more correctly it depends on the users' ability to choose "strong" passwords). Once the outsider succeeds to rlogin on the Dual-Homed Host machine he/she can access the entire Local Network.
- Dual-Homed Host runs proxy program for each service you want to permit, thus there is no more need for users to rlogin to the machine in order to access the Internet. They can communicate via proxy software.

The only host that can be accessed and thus attacked from the Internet is the Dual-Homed host machine. Thus it must have much greater level of security than the ordinary host on the Local Network. The excessive logging and auditing of

system state must be performed, only secure software and necessary software installed and so on.

This architecture is much more secure than the Screening Router Architecture. But still once the Dual-Homed Host is subverted the entire Local Network is vulnerable to attack.

5.5.5.3 Screened Host Architecture

This architecture consists of the Screening Router and Screened Host. Figure 5.7 shows this architecture.

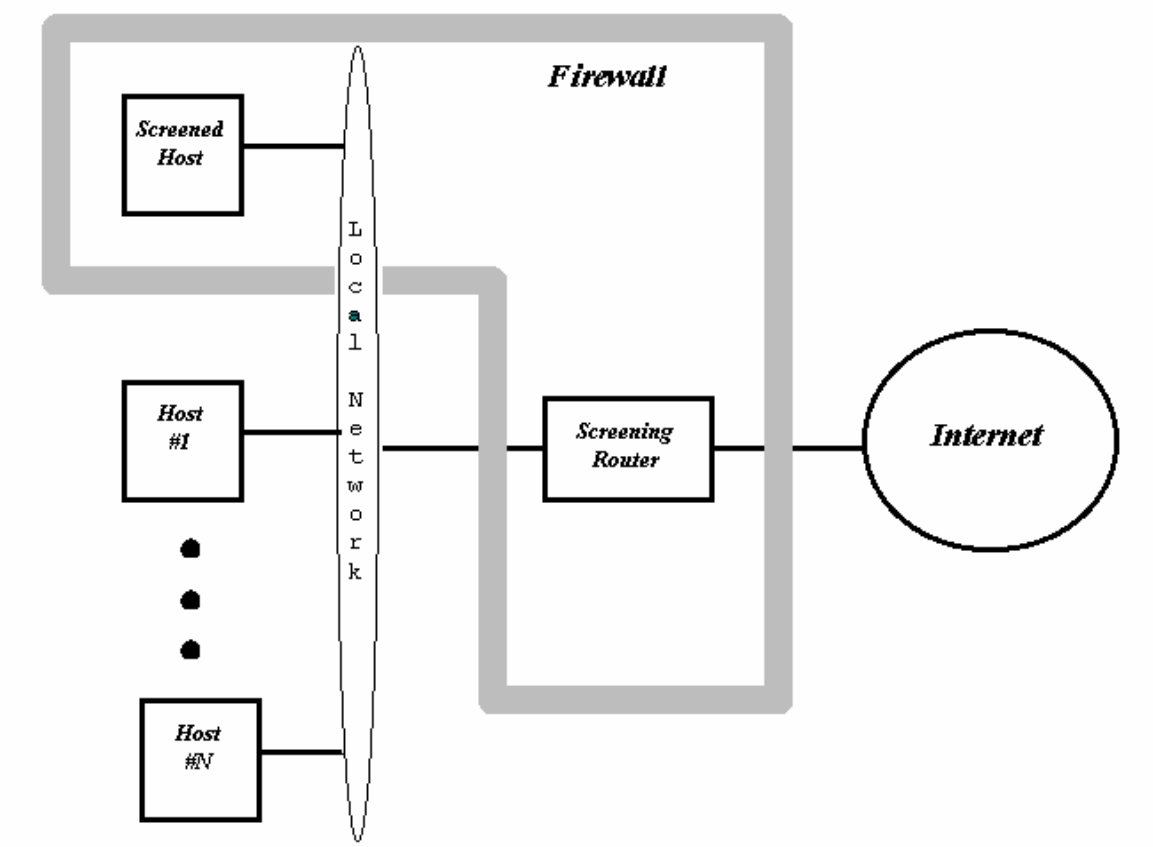


Figure 5.7

Screening Router is placed between the Local Network and the Internet and its role is to block all the traffic between those two networks but the one that originates on the Internet and goes to the Screened Host or the one that originates on the Screened Host and destined for the Internet. That is Screening

Router stops all the attempts to setup direct communication between ordinary host on the Local Network and the host on the Internet.

Screened Host is the host on the Local Network. It is the only host on the Local Network that can be accessed from the Internet and usually will run proxy programs for the allowed services. The other hosts on the Local Network must communicate with the Internet through proxy servers located on the Screened Host.

This architecture is more flexible than that of Dual-Homed Host with proxy services, because some secure services for which proxy software does not exist can be allowed to pass through Screening Router directly to a host on the Local Network.

Screened Host is also the only host that is subject to attack on an initial attempt. Thus an extra attention is paid to its security (because of this fact it is sometimes called in the literature “Bastion Host”). Once the Screened Host is subverted the attackers have access to all the hosts on the Local Network.

5.5.5.4 Screened Subnet Architecture

This architecture consists of the Screening Routers and Screened Hosts combined in such a way that when one of Screened Hosts is subverted the Local Network is not automatically open for an attack. In the figure bellow the Screened Subnet Architecture is shown using two Screening Routers and one Screened Host. Figure 5.8 shows this architecture.

5.5.5.5 Differences from the Screened Host Architecture

Screened Host is placed on the different physical segment than other hosts on the Local Network. Suppose that Screened Host is subverted. If it was connected to the same physical segment as other hosts in many network technologies it could to sniff all the traffic passing on the segment.

Local Network is guarded from the Screened Host by additional Screening Router. Thus in order to attack Local Network the attacker must pass through this additional router.

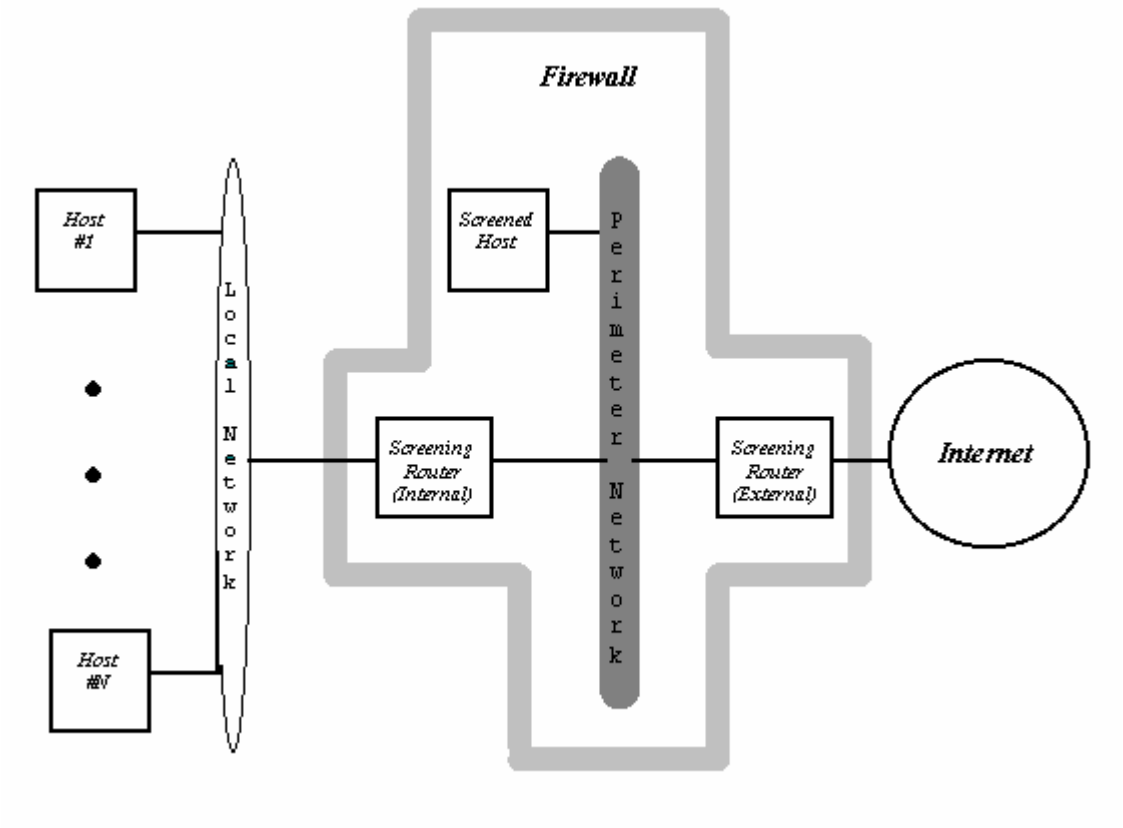


Figure 5.8

5.6 Dangers on the Internet

It is often the case that users are not prepared to deal with the issue of security or to protect themselves against security risks such as viruses or dialers. Many users only react when the damage is done, to the system or their finances. Surveys show that many users seriously deal with the issue of security when it is too late - when a virus or other damaging applications have emerged.

5.6.1 Dialers

One of the biggest threats for users who are directly connected with the Internet through a telephone connection such as ISDN or for modem users is dialers. These are dial-up programs, which use a phone line to create a connection with another system and in the past used expensive 0190 numbers.

These have become now 0900 numbers after a change in legislation. Nevertheless many PC users still incur huge financial losses due to these dialers. One of the most expensive dialers ever, called "Whirlpool", cost 900 Euros per dial-up to the Internet. Others cost approximately 80 Euros per minute. Since the change in legislation at the beginning of 2004, a maximum of 30 Euros per hour is permitted, and by law it must be disconnected after 60 minutes.

5.6.2 Viruses and worms

A virus is a program that nestles in other files and infects these. The infected file acts as a “host” through which the virus is further spread. While the computer is being infected the virus can destroy or manipulate existing data in the file and render them unusable. This can cause a loss of data.

Worms on the other hand do not require a host in order to be spread. They are independent programs which can be copied to other computers through security gaps for instance. However, they generally reach your computer by e-mail attachment. In order to activate the worm, the user simply has to run the attachment.

As a matter of fact, recently the number of worms has increased for which the user does not even need to do this in order to activate the worm. Some worms use security gaps in mail programs such as Outlook, through which the worm can be activated by reading a mail. After successfully infecting the computer, the worms are automatically sent to all e-mail addresses saved in the address book.

Pests also nestle in many Office products. You should be particularly skeptical when working with documents that contain macros. Macro is a programming language which controls and interprets files. The macro virus exploits exactly this option. It infects other files by executing its code. It mainly affects Excel tables and Word documents.

However, you should still be wary of tempting offers on Web pages which can only be accessed by using specific software.

5.6.3 Trojan horses

These are programs which give hackers access to external computers. Such programs pose as useful software or games for the users, in order to install a dangerous pest in the system, in the background… Now the hacker has the full control of your computer. They can access, manipulate or even delete data. Other Trojans carry out attacks on other computers so your PC is involved in such activities without you knowing about it.

5.6.4 Spyware

Spyware is a relatively new type of pest program. These programs collect information about the computer and its users and re-transmit information about the online behavior of the user to the author of the program without informing the PC user about these processes.

Most programs are circulated by online companies, in order to carry out targeted advertising for their products. Frequently these programs sneak into the system through incorrectly configured browsers. Others are installed on the computer through software such as music exchange marketplaces.

However, hackers have also discovered the benefits of these programs for themselves. They are less interested in the online behavior of the user and more in the keyboard entries. Keyloggers log all keyboard entries and send these back to the hackers. Passwords, access data for bank accounts, private correspondences - everything is recorded without the PC user noticing anything. After a certain period, the recorded information is simply sent by e-mail to a mailbox set up by the hacker.

5.7 IPSec

IPSec stands for Internet Protocol Security. IPSec is not a product in itself, but simply a set of protocols that govern the secure, private exchange of data across public networks, such as the Internet. It was developed by the Internet Engineering Task Force (IETF), and explained primarily in request for Comment’s (RFC’s) 2401-2412.

IPSec provides these security services at the IP layer; it uses IKE (Internet key exchange) to handle negotiation of protocols and algorithms based on local policy and to generate the encryption and authentication keys to be used by IPSec. IPSec can be used to protect one or more data flows between a pair of hosts, between a pair of security gateways, or between a security gateway and a host. IPSec works on Layer 3, the Network layer of the OSI model. By running on Layer 3, IPSec is able to function transparently to applications running on Layer 7; it is an extension to the standard IP protocols. IPSec is a framework of open standards that is used to create tunnels for Virtual Private Networks (VPN), and provide confidentiality, authenticity, and integrity of data through use of encryption algorithms. Combined with Internet Key Exchange (IKE), IPSec provides authentication and encryption services to protect unauthorized viewing or modification of data within your network or as it is transferred over an unprotected network, such as the public Internet.

There are two modes of IPsec operation: transport mode and tunnel mode.

In transport mode only the payload (message) of the IP packet is encrypted. It is fully-routable since the IP header is sent as plain text; however, it can not cross NAT interfaces, as this will invalidate its hash value. Transport mode is used for host-to-host communications.

In tunnel mode, the entire IP packet is encrypted. It must then be encapsulated into a new IP packet for routing to work. Tunnel mode is used for network-to-network communications (secure tunnels between routers) or host-to-network and host-to-host communications over the Internet.

IPsec is an obligatory part of IPv6, and is optional for use with IPv4. While the standard is designed to be indifferent to IP versions, current widespread deployment and experience concerns IPv4 implementations.

5.7.1 Different Phases

IPSec operates in two phases to allow the confidential exchange of a shared secret:

Phase 1: This phase handles the negotiation of security parameters required to establish a secure channel between two IPSec peers. Phase 1 is generally

implemented through the Internet Key Exchange (IKE) protocol. If the remote IPSec peer cannot perform IKE, we can use manual configuration with pre-shared keys to complete Phase 1.

Phase 2: This phase uses the secure tunnel established in Phase 1 to exchange the security parameters required to actually transmit user data.

5.7.2 Associated Protocols

Different Protocols associated with IPSec are:

- Authentication Header(AH)
- Encapsulated Security Payload(ESP)
- Internet Key Exchange(IKE)

5.7.3 Authentication Header

Authentication Header (AH) is intended to guarantee connectionless integrity and data origin authentication of IP datagrams. Further, it can optionally protect against replay attacks by using the sliding window technique and discarding old packets. AH protects the IP payload and all header fields of an IP datagram except for mutable fields, i.e. those that might be altered in transit. Mutable, therefore unauthenticated, IP header fields include TOS, Flags, Fragment Offset, TTL and Header Checksum. AH operates directly on top of IP using IP protocol number 51.

0	1	2	3
0 1 2 3 4 5 6 7	0 1 2 3 4 5 6 7	0 1 2 3 4 5 6 7	0 1 2 3 4 5 6 7
Next Header	Payload Length	RESERVED	

Security Parameters Index (SPI)
Sequence Number
Authentication Data (variable)

Figure 5.9 (An AH packet diagram)

An AH packet diagram is shown in figure 5.9. Here Fields meanings are-
Next Header - Identifies the protocol of the transferred data.
Payload Length - Size of AH packet.
Reserved - Reserved for future use (all zero until then).
Security Parameters Index (SPI) - Identifies the security parameters in combination with IP address.
Sequence Number - A monotonically increasing number, used to prevent replay attacks.
Authentication Data - Contains the data necessary to authenticate the packet

5.7.4 Encapsulating Security Payload

The Encapsulating Security Payload (ESP) extension header provides origin authenticity, integrity, and confidentiality protection of a packet. ESP also supports encryption-only and authentication-only configurations, but using encryption without authentication is strongly discouraged. Unlike the AH header, the IP packet header is not accounted for. ESP operates directly on top of IP using IP protocol number 50.

0								1								2								3							
0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7	0	1	2	3	4	5	6	7

<table>
<tr><td colspan="4">Security Parameters Index (SPI)</td></tr>
<tr><td colspan="4">Sequence Number</td></tr>
<tr><td colspan="4">Payload * (variable)</td></tr>
<tr><td></td><td colspan="3">Padding (0-255 bytes)</td></tr>
<tr><td></td><td></td><td>Pad Length</td><td>Next Header</td></tr>
<tr><td colspan="4">Authentication Data (variable)</td></tr>
</table>

Figure 5.10 An ESP Packet Diagram

Figure 5.10 shows the packet diagram of ESP. Here Field meanings are as:

Security Parameters Index (SPI) - Identifies the security parameters in combination with IP address

Sequence Number - A monotonically increasing number, used to prevent replay attacks.

Payload Data - The data to be transferred.

Padding - Used with some block ciphers to pad the data to the full length of a block.

Pad Length - Size of padding in bytes.

Next Header - Identifies the protocol of the transferred data.

Authentication Data - Contains the data used to authenticate the packet.

5.7.5 Internet key exchange

Internet key exchange (IKE) is the protocol used to set up a Security Association in the IPsec protocol suite. IKE uses a Diffie-Hellman key exchange to set up a shared session secret, from which cryptographic keys are derived. Public key techniques or, alternatively, a Pre-shared key (aka preshared secret), is used to mutually authenticate the communicating parties. IKE builds upon the Oakley protocol.

5.7.5.1 Architecture

Most IPsec implementations consist of an IKE daemon that runs in user space and an IPsec stack in the kernel that processes the actual IP packets.
User space daemons have easy access to mass storage which contains configuration information such as the IPsec endpoint addresses, keys and certificates as required. Kernel modules on the other hand can process packets efficiently and with minimum overhead - which is important for performance reasons.
The IKE protocol uses UDP packets, usually on port 500 and generally requires 4-6 packets with 2-3 turn-around times to create an SA on both sides. The negotiated key material - say an AES key and endpoint information (which IP endpoints and ports we are protecting) as well as what type of IPsec tunnel has been created - is then given to the IPsec stack. It in turn intercepts the relevant IP packets if and where appropriate and performs encryption/decryption as required. Implementations vary on how the interception of the packets is done. Some use virtual devices, others take a slice out of the firewall - it varies

5.8 The Future of IP (IPv6)

One of the newest major standards on the horizon is IPv6. Although IPv6 has not officially become a standard, it is worth some overview. It is very possible that this information will change as we move closer to IPv6 as a standard, so you should use this as a guide into IPv6, not the definitive information.
A number of books are now being published that cover in detail this emerging standard. All the RFCs available on the Internet have the raw details on how this standard is developing. However, these documents are difficult to interpret at first glance and require some commitment to going through any number of RFCs pertaining to many subjects all related to IPv6 development.
Internet Protocol Version 4 is the most popular protocol in use today, although there are some questions about its capability to serve the Internet community much longer. IPv4 was finished in the 1970s and has started to show its age. The main issue surrounding IPv6 is addressing—or, the lack of addressing—because

many experts believe that we are nearly out of the four billion addresses available in IPv4. Although this seems like a very large number of addresses, multiple large blocks are given to government agencies and large organizations. IPv6 could be the solution to many problems, but it is still not fully developed and is not a standard—yet.

Many of the finest developers and engineering minds have been working on IPv6 since the early 1990s. Hundreds of RFCs have been written and have detailed some major areas, including expanded addressing, simplified header format, flow labeling, authentication, and privacy.

Expanded addressing moves us from 32-bit address to a 128-bit addressing method. It also provides newer unicast and broadcasting methods, injects hexadecimal into the IP address, and moves from using "." to using ":" as delimiters. Figure 5.11 shows the IPv6 packet header format.

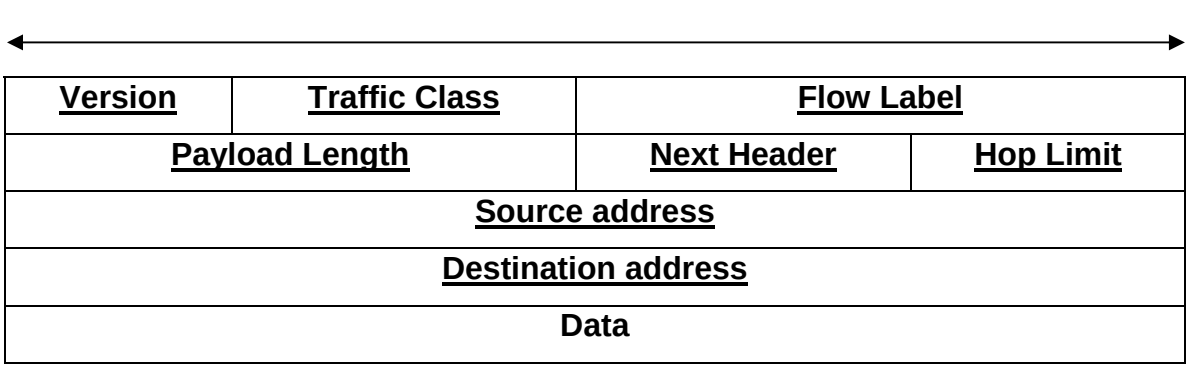


Figure 5.11 (IPv6 Packet Header Format)

The simplified header is 40 bits long and the format consists of Version, Class, Flow Label, Payload Length, Next Header, Hop Limit, Source Address, Destination Address, Data, and Payload fields.

Version (4 bits) - IPv6 version number.

Traffic Class (8 bits.) - Internet traffic priority delivery value.

Flow Label (20 bits) - Used for specifying special router handling from source to destination(s) for a sequence of packets.

Payload Length (16 bits unsigned) - Specifies the length of the data in the packet. When cleared to zero, the option is a hop-by-hop Jumbo payload.

Next Header (8 bits) - Specifies the next encapsulated protocol. The values are compatible with those specified for the IPv4 protocol field.

Hop Limit (8 bits unsigned) - For each router that forwards the packet, the hop limit is decremented by 1. When the hop limit field reaches zero, the packet is discarded. This replaces the TTL field in the IPv4 header that was originally intended to be used as a time based hop limit.

Source address (16 bytes) - The IPv6 address of the sending node.

Destination address (16 bytes) - The IPv6 address of the destination node.

Some of the benefits of IPv6 seem obvious: greater addressing space, QoS(Quality of Service), and better routing performance and services. However, a number of barriers must be overcome before the implementation of IPv6. The biggest question for most of us will be what the business need is for moving from current IPv4 to IPv6. The killer app has not appeared yet, but it may be closer than we think. The second consideration is the cost—it may not have much to do with hardware replacement cost. All the larger routers have upgradable OSs IOS; the only necessity is the commitment to upgrading IOS. More likely to do with training and support of minor IP devices such as printers and network faxes, they will support the new address space. IPv6 has schemes to support old and new, however, so this may not even be a barrier. The last issue to consider is training: This will need to happen sooner or later because we all need to start thinking about 128-bit addressing based on MAC addresses in HEX. This involves all new ways of addressing and will be an uncomfortable change for many people.

This conclusion may seem negative, but the greater good will overpower all the up-front issues. The issue is not whether you will have to move to IPv6, but when! We all need IPv6; the increased address space is needed for the growth of IP appliances. IP-ready cars are already shipping today. This requires mobility, which is addressed in IPv6.

5.9 Internet Tools

There are number of tools available online to perform various functions on internet. Some well-known tools are E-mail, Search Engines, Chat, Instant Messengers, Discussion forms, FTP, and Telnet.

5.9.1 E-mail

Electronic mail (e-mail) is one of the most widely used services on the Internet. E-mail is easy to send, read, reply to, and manage. E-mail is fast and convenient. For these reasons, e-mail has grown from a simple service offered to researchers for communicating ideas and results into a complex, talented messaging system.

- It is much easier to write an e-mail message than to write a formal paper letter or note. Many studies have shown that recipients are much more likely reply to an e-mail message than a written request, primarily because of the ease of formulating the response.
- E-mail is very fast, the mail usually reaches its destination in a matter of seconds
- E-mail is also economic. It is much cheaper to send an email message than a letter or to make a long distance telephone call.
- You can send letters, notes, files, data or reports using the same technique
- It is asynchronous, i.e. you don't have to be present when your mail arrives. You can receive and read it at any convenient time.
- It is possible now to read and write e-mails on portable devices such as mobile phone. So you can remain in touch even when you are traveling.

5.9.2 Search Engines

Search Engines play very important role in browsing as it is not possible to remember the addresses of all the web sites. Moreover they find the web pages of interest immediately. We only have to enter the worlds related to particular interest and search engines display the web site containing those worlds. Google

(http://google.com), Yahoo(http://www.yahoo.com), MSN(http://www.msn.com) and Khoj (http://www.khoj.com) are the popular search engines.

5.9.3 Chat

Chatting is one of the popular tools among the young generation. One can chat with a person who may be on the other corner of world and exchange information. The advantages of using “chat” are-

- It is cheap.
- Response is immediate.
- One can select person for chatting depending upon the interest or region.
- It is also possible to exchange voice and video messages.

Some of the popular and free chat clients are Yahoo messenger, Rediff bol, and msn messenger.

5.9.4 Discussion Forms

Discussion forms are platforms to post your messages or queries, get knowledge from, share your experiences and exchange your views. Discussion forms play great roles when you want to know some thing from real persons rather than static information available on line. There are various forms on technical, political, sports, movies and medical issues where you can discuss your problems.

5.9.5 FTP

One of the reasons for creating the Internet was that people could exchange ideas and results of their work. FTP, short for File Transfer Protocol is used for transferring files over the Internet between a remote computer and your computer. FTP provides for access control and negotiation of file parameters. FTP provides for access control and negotiation of file parameters. FTP is most effective when you know the exact location -file name, directory name and Internet name of the remote computer where the file is located. FTP allows you to connect with an FTP server, browse through its directories and files, and then retrieve a copy of any file useful to you.

5.9.6 Telnet

The Internet is a system of networked computers. Networks allow users from one computer to access information or run programs on another computer on the network. In the Internet environment this is possible through a tool called the Telnet. (This tool allows a user at a remote site to access facilities, software or data at another site (remote site). While using telnet it is as though you are directly connected to the remote site. As with most UNIX systems it is important to have a valid username and password for logging onto these remote sites, though most sites allow anyone to log onto their systems.

5.10 Summary

Internet, being on public domain requires many security aspects. Encryption may be useful as messages passes through various channels. Firewall controls the traffic passed to and fro from LAN to internet, thus can be affectively employed to stop the misuse of net with in an organization and make the LAN secure from outside dangers. The future of IP is IPv6 as it uses 128 bit addressing method and better routing performance.

Freely available internet tools such as e-mail, search engines, chatting and discussion forms have made our life simpler and interesting.

5.11 Self Assessment Questions

Q1. What are the measures implemented on internet to make it secure?

Q2. Which firewall design will you prefer for your institute?

Q3. What are the difficulties to implement IPv6?

Q4. How internet tools have made our life simpler?

Q5. Why security on internet is big issue? What are the different security concerns on internet?

Q6. Explain the role of firewall on internet.

5.12 References

1. Firewall, RL Armstrong, 2005
2. Cisco Router Firewall Security, Richard A. Deal, Cisco Press, 2004
3. How to Cheat at Managing Information Security, Mark Osborne, Syngress Publishing, 2006
4. Computer Viruses: From Theory to Applications, Eric Filiol, Springer, 2005

Paper Code: MCA- 405	**Author: Parvinder Singh**
Paper Name: Computer Networks-II	**Vetter: Dinesh Kumar**
Lesson Number: 06	

High Speed Network

Structure

6.0 Objective

The objective of this chapter is to cover various high speed networking technologies such as frame relay, ATM and high speed LANs such as Fast Ethernet, and Gigabit Ethernet, Fibre Channel and high speed wireless LAN.

6.1 Introduction

The term “high-speed network” is a relative name. It seems not long ago (1970) that 4,800 bits per second leased line was considered very high in speed. In the 1990s, 2 Mbps wide area links and LANs using speeds of 10 and 16 Mbps had become universal. Today very much higher speeds (hundreds of megabits per second) are common. The networking techniques and technologies currently in use are unable to operate efficiently at the newly available higher speeds. This chapter describes the new approaches that are required to enable efficient use of the new high speeds. So rather than defining “high speed” to mean any particular speed, for the purposes of this chapter “high speed” is held to mean “any speed that requires the use of new networking techniques for efficient operation”. In practice this dividing line is somewhere around 100 Mbps for LANs and about 35 Mbps for wide area communications.

6.2 Need for High Speed Networks

It is always considered to have a very good speed in communication. Generally high speed in network is required due to following reasons-

1. To implement new data applications using graphics and image processing, etc.
2. To do existing applications in better ways. For example instead of using coded commands and responses (such as are typical of banking and airline applications), users would like fast response full-screen applications so that staff do not need the specialized skills that previous systems demanded.
3. To rationalize the many dissimilar networks that major users have. Many users have SNA, DECnet, TCP/IP and X.25 data networks as well as logical networks of FAX machines and a voice network. For management reasons (and for cost

optimization, though this is becoming less and less of a factor) users want to have only one network to handle all traffic.

4. Integration of voice, video, and data. Most large users have a voice network and a separate data network, and for cost and manageability reasons would like to integrate these. In addition there is the future possibility of voice and data integrated applications. Many users see video as an important opportunity and would like to be able to handle this traffic.

6.3 Performance Attributes

6.3.1 Bandwidth Cost

The cost of transmitting “x” bits per second continues to reduce exponentially.

6.3.2 Attachment Cost

The cost of connecting a link (be it copper wire or optical fiber) from a user’s location to the public network continues to increase with inflation. Having any connection at all involves “digging up the road” and this cost continues to rise. Today’s technology enables us to use a single connection for many different things simultaneously and to use significantly higher transmission rates, but the cost of having a single connection is increasing.

6.3.3 Error Rates

Error rates on communication links have improved dramatically. On an analog telephone line with a modem, a typical rate might have been 10^{-5} - or one error in every 100,000 bits transmitted. On a fiber connection this rate may be as low as 10^{-11} - a million times better.

6.3.4 Error Characteristics

The kinds of errors have changed also. On a digital connection errors tend to occur in bursts, whereas on an analog line they usually occur as single or double bit errors.

6.3.5 Propagation Delay

At 80% of the speed of light, the speed of message propagation on a communication line has not changed.

6.3.6 Storage Effect

At very high speeds long communication lines store a large amount of data. For example, a link from New Delhi to Chennai involves a delay of roughly 20 milliseconds. At the 100 Mbps speed of FDDI this means that two million bits are in transit at any time in each direction. So the link has approximately half a million bytes stored in transit. This has a critical effect on the efficiency of most current link protocols.

6.3.7 Computer Technology

Computers continue to become faster and lower in cost at an approximate rate of 30% per annum compounded. However, this is not as simple as it sounds - some things (for example the cost of main storage) have reduced faster than others (for example the costs of power supplies, screens, and keyboards).

Data links are now considerably faster than most computers that attach to them. In the past, the communications line was the limiting factor and most computer devices were easily able to load the link at perhaps 95% of its capacity. Today, very few computer devices are capable of sustaining a continuous transfer rate of anything like the speed of a fast fiber link. The “state of the art” in public network fiber facilities today is a transmission rate of 2.4 Gbps, but rates of many times this are functioning in laboratories.

6.4 Traffic Characteristics

Many organizations see the new lower communication cost structure as an opportunity for:

1. Doing old applications better.
2. Doing new applications that were not feasible (or indeed imaginable) before.

The first requirement is to integrate existing networks into a single network. The motivation for this is not only to save money on links but to provide a better networking service by integrating the many disparate networks into a single coherently managed unit.

The kinds of existing networks that users want to integrate can be summarized as follows:

- Traditional data networks

- Voice networks
- Interconnected LAN networks
- Multiprotocol networks

In addition there are opportunities for applications using:

- Image
- Full-motion video

Traditional data networks were built to handle both interactive and batch data but were not built to handle image, voice or video traffic. The new types of traffic put a completely new set of requirements onto the network.

If the user requirements outlined above (integration of data, voice, video, image, etc.) are to be satisfied by packet networks, then clearly a new type of packet network will be needed. Network nodes will need to handle the full data throughput capacity of the new high-speed links (one million packets per second - plus) and network architectures will need to accommodate the unique characteristics of voice and video traffic.

The requirements may be summarized as follows:

6.4.1 Very High Node Throughput

Nodes must be able to route (switch) data at the peak combined rate of all links connected to them. In corporate networks this might mean a maximum of perhaps 20 links at 155 Mbps, but this seems a little high for the decade of the 1990s. More likely would be a switch with less than 20 links where perhaps four of them are 155 Mbps and the rest might be at the “T3” rate of 45 Mbps.

But corporate private networks are one thing. Public telecommunications networks are something else. The proposal with ATM (B_ISDN) is that packet (cell) switching should become the basis of a multi-function network, which will replace the world’s telephone network. To do this, a mainline trunk exchange (probably a cluster of switching nodes) would need to handle perhaps 100 links of 620 Mbps today and perhaps the same 100 links would be running at 2.4 Gbps by the time the system was built.

Using 53-byte cells, a 2.4 Gbps link can carry just less than six million cells per second in each direction. The example is a little extreme but the principle is clear.

We are going to need the ability to process cells at rates of well above one hundred million per second for Broadband ISDN to become a reality.

6.4.2 Minimal Network Transit Time

This is a critical requirement for voice and it depends on the structure of the network but should be less than one millisecond per node traversed plus propagation delay at about 5.5 msec per kilometer.

6.4.3 Minimal Variation in Network Transit Time

When any traffic with a constant bit rate at origin and destination travels through a network, the variations in network delay mean that a buffer somewhat larger than the largest foreseeable variation in transit time is needed. This buffer introduces a delay and for practical purposes can be considered a net addition to network transit time.

To meet the above requirements networks will need to have the following characteristics:

6.4.4 Totally Hardware Controlled Switching

There is no way that current software-based packet switched architectures can come to even one hundredth of the required throughput - even assuming much faster processors.

However, there are several hardware switching designs that will meet the required speeds at (predicted) reasonable cost.

6.4.5 Suitable Network Architecture

The network architecture must make it possible for the data switching component in a node to decide the destination to which an incoming packet should be routed at full operating speed. The network architecture must provide mechanisms for the stable operation and management of the network but the data switching element must not need to get involved in extraneous protocols.

6.4.6 Link Error Recovery

Recovery from transient link errors by retransmission (for voice traffic), as is usual for data traffic, can seriously conflict with the requirement for uniform delivery rates. For voice, a delayed packet is worse than a packet in error. However, by the nature of packetization, it is necessary that packets contain a

header which carries routing information (identification) so the destination switch can route it to the appropriate destination. An error in this information will cause a packet to be routed to the wrong destination and a packet to be lost from the correct circuit. But these very high-speed networks are planned to operate solely over digital (preferably fiber optical) circuits. Error rates on these circuits are around ten thousand times better than they were for traditional analog data links.

For the data portion of the packet or cell, error checking and recovery can be applied on an end-to-end basis especially if the error rates experienced on links is very low. The header portion is not so fortunate. An error in the header can cause a packet to be misrouted to the wrong destination. The network must at least check the headers.

6.4.7 Packet Length

Short (less than 64 bytes), fixed-length packets or cells are an attractive option because:

1. Their fixed-length nature gives a uniform transmission time (per cell) characteristic to the queueing within a node for an outbound link. This leads to a more uniform transit-time characteristic for the whole network.
2. The shorter the cell the shorter the time needed to assemble it and hence the shorter the delay characteristic for voice.
3. Short, fixed-length cells are easy to transfer over a fixed-width processor bus, and buffering in link queues is a lot easier and requires less processor logic.

One elegant solution to both the network delay and error recovery problems would be to use very short packets (perhaps 32 bytes) of fixed length. If this is done then Error Correcting Codes (ECC) can be used as a recovery from transient link errors. Two bytes of ECC are required for every 8 bytes of data. A 32-byte packet would then have a routing header (2 or 4 bytes) included within it and one or four ECC 2-byte groups appended to it (one if it is thought necessary only to check the header, two if the data is to be error recovered also). Therefore, a packet would be either 34 or 40 bytes. (This represents an overhead on the transmission channel in the full ECC case of 20%.) It happens that the use of ECC in this way for a voice packet is considered wasteful and unnecessary. The

loss of a packet or two (provided it is relatively infrequent) or the corruption of a few bits of data is not considered to be significant.

The international standard for cell size is now 48 bytes (for ATM). In ATM the header is checked for validity but the data within the cell is not (or, rather, that checking and error recovery on the data within a frame (group of cells) is left to the end-to-end protocol called the “adaptation layer”).

However, there is another side. Video transmission is fine with packet sizes of over a thousand bytes. Data transmission can be achieved with low overhead if the packet size adopted is large enough to carry the largest natural data block as produced by the user’s application.

The longer the packet the fewer packets per second must be switched for a given data throughput.

6.4.8 Flow Control

Control of congestion is a critical matter in any packet switching environment. Traditional techniques of flow control are not possible at very high packet rates because they require significant amounts of programmed logic to operate on every packet.

In a high-speed switch, input rate regulation and capacity reservation are the appropriate techniques. These can be agreed by the control processors when a connection is started and enforced at the entry points of the network.

Congestion Control

Congestion occurs when a node builds up too much data for its internal buffers to process. This can happen even in data networks with very detailed explicit flow controls.

One way to handle congestion is to avoid it. Good flow controls can help in avoiding congestion. Another sure way of handling congestion is to make sure that the maximum demand that can ever be placed on the network can be met at all times. This means running links and nodes at average utilizations of around 10 or 20% at the peak. But this foregoes the benefits of sharing the network. If the network is to process variable rate data (say voice) from many thousands of

users simultaneously, and if no single user can make a peak demand sufficient to be noticed, then the statistics of the situation work for us.

Congestion becomes a problem where there are a number of sources that can individually place a significant demand on the network (such as in variable-rate video). In this case a small number of users (as few as 10 perhaps) might be able to each make peak demands simultaneously and bring the whole network to a standstill. The trick here is to avoid the situation where any single user can make a significant demand on the network.

But some types of traffic change radically over time. Data traffic peaks at different times in a business day. Batch data peaks during the night. When congestion occurs packets must be discarded. For some data types (voice, video) coding can be such that low priority packets can be discarded with the net effect of a “graceful degradation” of the service. If these packets are marked as discardable in some way (this is a feature of both ATM and Paris), then the system can alleviate congestion by discarding these.

If congestion becomes very serious, then the network will need to discard packets not marked as discard able. The network should have a way of prioritizing traffic by service class so that an intelligent packet discard strategy may be adopted.

This packet discard strategy must be performed by the (hardware) data switching element. The discard strategy must be very simple.

6.4.9 Sequential Delivery

If packets applying to one conversation are allowed to take different routes through the network (for load balancing for example) then they must be resequenced before delivery to the receiver. However, this means that each would have to carry a sequence number (more overhead) and the technique would result in “bursty” uneven delivery.

To overcome this, delivery would then need to be buffered sufficiently to even out the bursts. This would add cost but more importantly it would add to the transit delay and thus degrade the quality.

In a high-speed network this means that each connection must be limited to a fixed path through the network.

6.4.10 Priorities

There is no consensus yet on whether transmission priorities are relevant in a high-speed network. A transmission priority may be given to a packet and that priority enables it to “jump the queue” ahead of lower priority packets when being queued for transmission within a node.

Within a tightly controlled traditional packet networking system such as SNA, the system of priorities has worked well. It gives better response time to higher priority traffic and also enables the use of much higher resource (link and node) loadings than would be possible without them.

But at such high speed, with relatively small cells (at the speeds we are considering even a 4KB block is small - in time), many people suggest that the cost of implementing priorities may be greater than it is worth.

Most studies of high-speed node technology suggest that the total switching (processing, queueing and transmission) in the kind of node under discussion will be much less than one millisecond.

Other kinds of priority are, however, considered essential. In a large network there needs to be some control and prioritization of the selection of routes through a network, depending on the required service characteristics for a particular class of service.

In addition it seems generally agreed that a service class type of priority should be used to decide which packets to discard at times of network congestion.

6.4.11 End-to-End Protocols and “Adaptation”

The characteristics of a high-speed network developed thus far are such that it gives very high throughput of very short packets, but in the case of congestion or of link errors packets are discarded. To provide a stable service, the network needs to have processing at the entry and exit points of the network. This processing will, for example, break long frames of data up into cells and reassemble at the other end.

In addition, for data traffic it should implement a Frame Check Sequence (FCS) calculation to identify frames containing errors. It may also have a retransmission protocol to recover from data errors and lost packets, etc. (Or it may just signal to the user that there has been a problem and allow the user to do recovery.) Each type of network traffic requires different adaptation layer processing.

6.5 Network Backbone

6.5.1 Frame Relay

Frame Relay is a high-performance WAN protocol that operates at the physical and data link layers of the OSI reference model. Frame Relay originally was designed for use across Integrated Services Digital Network (ISDN) interfaces. Today, it is used over a variety of other network interfaces as well. Frame Relay is an example of a packet-switched technology. Packet-switched networks enable end stations to dynamically share the network medium and the available bandwidth. The following two techniques are used in packet-switching technology:

- Variable-length packets
- Statistical multiplexing

Variable-length packets are used for more efficient and flexible data transfers. These packets are switched between the various segments in the network until the destination is reached.

Statistical multiplexing is a type of communication link sharing. In statistical multiplexing, a fixed bandwidth communication channel is divided into several variable bit-rate digital channels. The link sharing is adapted to the instantaneous traffic demands of the data streams that are transferred over each channel. This is an alternative to creating a fixed sharing of a link, such as in general time division multiplexing and frequency division multiplexing. When performed correctly, statistical multiplexing can provide a link utilization improvement, denoting the statistical multiplexing gain.

Statistical multiplexing techniques control network access in a packet-switched network. The advantage of this technique is that it accommodates more flexibility and more efficient use of bandwidth. Most of today's popular LANs, such as Ethernet and Token Ring, are packet-switched networks.

Frame Relay often is described as a streamlined version of X.25, offering fewer of the robust capabilities, such as windowing and retransmission of last data that are offered in X.25. This is because Frame Relay typically operates over WAN facilities that offer more reliable connection services and a higher degree of reliability than the facilities available during the late 1970s and early 1980s that served as the common platforms for X.25 WANs. As mentioned earlier, Frame Relay is strictly a Layer 2 protocol suite, whereas X.25 provides services at Layer 3 (the network layer) as well. This enables Frame Relay to offer higher performance and greater transmission efficiency than X.25, and makes Frame Relay suitable for current WAN applications, such as LAN interconnection.

6.5.1.1 Frame Relay Devices

Devices attached to a Frame Relay WAN fall into the following two general categories:

- Data terminal equipment (DTE)
- Data circuit-terminating equipment (DCE)

DTEs generally are considered to be terminating equipment for a specific network and typically are located on the premises of a customer. In fact, they may be owned by the customer. Examples of DTE devices are terminals, personal computers, routers, and bridges.

DCEs are carrier-owned internetworking devices. The purpose of DCE equipment is to provide clocking and switching services in a network, which are the devices that actually transmit data through the WAN. In most cases, these are packet switches.

Figure 1 shows the relationship between the two categories of devices.

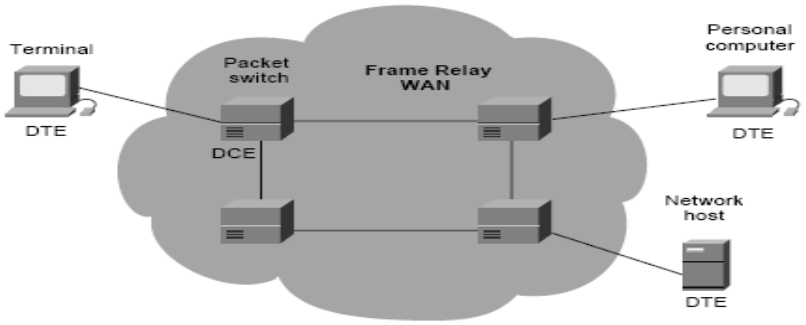


Figure 1 DCEs Generally Reside Within Carrier-Operated WANs

The connection between a DTE device and a DCE device consists of both a physical layer component and a link layer component. The physical component defines the mechanical, electrical, functional, and procedural specifications for the connection between the devices. One of the most commonly used physical layer interface specifications is the recommended standard RS-232 specification. The link layer component defines the protocol that establishes the connection between the DTE device, such as a router, and the DCE device, such as a switch. This chapter examines a commonly utilized protocol specification used in WAN networking: the Frame Relay protocol.

6.5.1.2 Frame Relay Network Implementation

A common private Frame Relay network implementation is to equip a T1 multiplexer with both Frame Relay and non-Frame Relay interfaces. Frame Relay traffic is forwarded out the Frame Relay interface and onto the data network. Non-Frame Relay traffic is forwarded to the appropriate application or service, such as a private branch exchange (PBX) for telephone service or to a video-teleconferencing application.

A typical Frame Relay network consists of a number of DTE devices, such as routers, connected to remote ports on multiplexer equipment via traditional point-to-point services such as T1, fractional T1, or 56-Kb circuits. An example of a simple Frame Relay network is shown in Figure 2

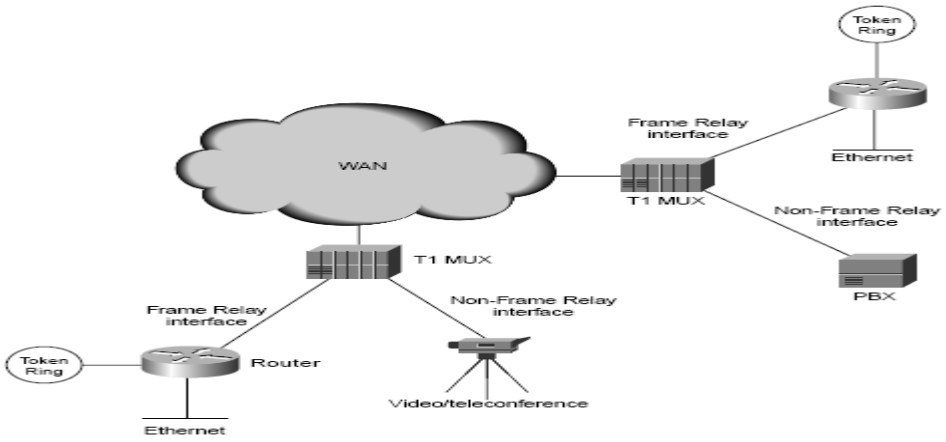


Figure 2 A Simple Frame Relay Network Connects Various Devices to Different Services over a WAN

The majority of Frame Relay networks deployed today are provisioned by service providers that intend to offer transmission services to customers. This is often referred to as a public Frame Relay service. Frame Relay is implemented in both public carrier-provided networks and in private enterprise networks. The following section examines the two methodologies for deploying Frame Relay.

6.5.1.3 Frame Relay Frame Format

To understand much of the functionality of Frame Relay, it is helpful to understand the structure of the Frame Relay frame. Figure 3 depicts the basic format of the Frame Relay frame.

Flags indicate the beginning and end of the frame. Three primary components make up the Frame Relay frame: the header and address area, the user-data portion, and the frame check sequence (FCS). The address area, which is 2 bytes in length, is comprised of 10 bits representing the actual circuit identifier and 6 bits of fields related to congestion management. This identifier commonly is referred to as the data-link connection identifier (DLCI). Each of these is discussed in the descriptions that follow.

Field length in bits

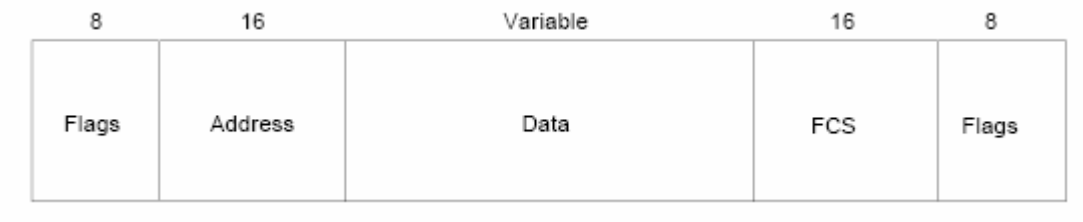


Figure 3 Five Fields Comprise the Frame Relay Frame

The following descriptions summarize the basic Frame Relay frame fields illustrated above.

Flags—Delimits the beginning and end of the frame. The value of this field is always the same and is represented either as the hexadecimal number 7E or as the binary number 01111110.

Address—Contains the following information:

–DLCI—The 10-bit DLCI is the essence of the Frame Relay header. This value represents the virtual connection between the DTE device and the switch. Each virtual connection that is multiplexed onto the physical channel will be represented by a unique DLCI. The DLCI values have local significance only, which means that they are unique only to the physical channel on which they reside. Therefore, devices at opposite ends of a connection can use different DLCI values to refer to the same virtual connection.

–Extended Address (EA)—The EA is used to indicate whether the byte in which the EA value is 1 is the last addressing field. If the value is 1, then the current byte is determined to be the last DLCI octet. Although current Frame Relay implementations all use a two-octet DLCI, this capability does allow longer DLCIs to be used in the future. The eighth bit of each byte of the Address field is used to indicate the EA.

–C/R—The C/R is the bit that follows the most significant DLCI byte in the Address field. The C/R bit is not currently defined.

–Congestion Control—This consists of the 3 bits that control the Frame Relay congestion-notification mechanisms. These are the FECN, BECN, and DE bits, which are the last 3 bits in the Address field.

Forward-explicit congestion notification (FECN) is a single-bit field that can be set to a value of 1 by a switch to indicate to an end DTE device, such as a router, that congestion was experienced in the direction of the frame transmission from source to destination. The primary benefit of the use of the FECN and BECN fields is the capability of higher-layer protocols to react intelligently to these congestion indicators. Today, DECnet and OSI are the only higher-layer protocols that implement these capabilities.

Backward-explicit congestion notification (BECN) is a single-bit field that, when set to a value of 1 by a switch, indicates that congestion was experienced in the network in the direction opposite of the frame transmission from source to destination.

Discard eligibility (DE) is set by the DTE device, such as a router, to indicate that the marked frame is of lesser importance relative to other frames being transmitted. Frames that are marked as "discard eligible" should be discarded before other frames in a congested network. This allows for a basic prioritization mechanism in Frame Relay networks.

Data—Contains encapsulated upper-layer data. Each frame in this variable-length field includes a user data or payload field that will vary in length up to 16,000 octets. This field serves to transport the higher-layer protocol packet (PDU) through a Frame Relay network.

Frame Check Sequence—Ensures the integrity of transmitted data. This value is computed by the source device and verified by the receiver to ensure integrity of transmission.

6.5.2 ATM

Asynchronous Transfer Mode (ATM) is a high-speed connection-oriented switching and multiplexing technology that allows for the execution of high speed telecommunications applications, such as- telemedicine, distance learning, data applications -- huge files, video conferencing, video broadcast and multimedia.

Asynchronous Transfer Mode (ATM) is a cell relay, network and data link layer protocol which encodes data traffic into small (53 bytes; 48 bytes of data and 5 bytes of header information) fixed-sized cells. This is instead of variable sized

packets (sometimes known as frames) as in packet-switched networks (such as the Internet Protocol or Ethernet).

6.5.2.1 Why is ATM Important?

An ATM backbone allows carriers and network providers to provide multiple services on a single ATM network, under a single network management system. As a result, voice, data, image, video and multimedia traffic can use the same network.

An ATM -based network is capable of serving new applications and improving network performance and productivity all at justifiable costs. For example, lives can be saved by extending medical expertise with ATM applications. In addition, ATM makes high quality tracking available to remote locations in an affordable manner. ATM also provides enhanced general information transfer.

6.5.2.2 Higher Network Availability

ATM switches are highly flexible because their switching is based on hardware, not software. This factor also allows very high speeds since minimal processing has to occur at each switch.

Since they provide a single architecture for all traffic, ATM networks will be simpler to manage than others.

6.5.2.3 High Bandwidth Scalability and Flexibility

The ATM switching fabric is inherently scalable and ATM supports a broad range of services so an ATM network put in place today can grow and change as network needs change.

6.5.2.4 Cost Benefits

ATM can provide many cost benefits through:

- An improved cost/performance ratio
- The consolidation of applications
- The better use of facilities with the combination of all services on a single switched backbone
- Compatibility with currently deployed physical networks
- Uniform technology
- Lower network management costs

- A longer life cycle, future-proof network.

6.5.2.5 Performance

- Flexible access speeds from T1 - OC3, OC12 - OC48
- Well defined connection procedures
- Higher aggregate bandwidth
- Dedicated bandwidth per connection
- Low delay per switching node

6.5.2.6 ATM Architecture

ATM is divided into layers (See Fig. 4).The physical layer is divided into two parts. The ATM physical medium sublayer is responsible for transmission of data over the physical medium, regardless of the type of medium used. ATM was originally designed to operate over fiber optics but because of the slow deployment of fiber, was later modified to operate over copper and coaxial facilities as well.

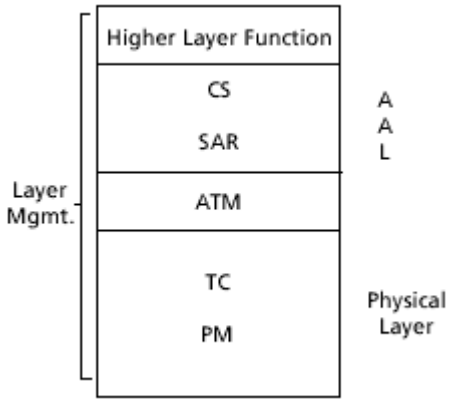


Figure 4 ATM Layers

The physical medium sublayer is responsible for receiving and transmitting bit streams in a continuous method. This is important to channelized services with rely on constant bit streams to maintain synchronization. When the bit stream stops, channelized equipment interprets the condition as an error and releases the virtual connection. bit synchronization is also maintained by this sublayer.

The transmission convergence sublayer is responsible for the transmission and reception of frames over framed transmission facility, such as T-3. ATM cells are packed into these frames and unpacked at the remote end. This sublayer also

performs error detection/correction but only on the ATM header. This prevents the cells from being sent to the wrong destination.

Cell rate decoupling is used when a continuous data stream is required at the physical layer, as in SONET and channelized facilities such as DS1. Cell rate decoupling sends special "idle" cells over the framed facility and discards any idle cells it receives. Idle cells are necessary to maintain a connection in channelized facilities because the channel bank equipment must always see a constant bit rate transmission, or it disconnects the channel. When nothing is being sent over a channelized facility, idle flags are transmitted (this is also used to maintain clock synchronization between two endpoints). Idle cells are not recognized by the ATM layer.

The functions of the transmission Convergence Sublayer (CS) differ depending on the medium being used. For instance, if SONET is the medium, the physical layer requires a different set of functions than a DS-3 medium would require. This sublayer provides whatever services are needed by each type of medium.

There are some specific functions required for DS3 and 100-Mbps interfaces. The ATM physical layer provides a convergence protocol (Physical Layer Convergence Protocol, PLCP), which maps ATM cells onto a DS3. The interface supports 44.736 Mbps. ATM cells are mapped into a DS3 PLCP data unit, which is then mapped into the DS3 payload. The DS3 PLCP is not aligned to the DS3 framing bits.

The 100-Mbps access was intended for private UNIs. Private UNIs are not as complex as public UNIs, which must provide higher reliability and complex monitoring. The specification is based on the FDDI physical layer.

The ATM layer is responsible for multiplexing cells over the interface. ATM must read the VPI/VCI of incoming cells, determine which link cells are to be transmitted over, and place new VPI/VCI values into the header. At endpoints, the ATM layer generates and interprets cell headers (endpoints do not route cells).

The ATM layer supports the following connection types:

Point-to-point Virtual Channel Connection (VCC)

Point-to-multipoint VCC

Point-to-point Virtual Path Connection (VPC)

Point-to-multipoint VPC

A VCC is a single connection between two endpoints. A VPC is a bundle (or group) of VCCs carried transparently between two end points.

The AAL is used mostly by endpoints. It is divided into two sublayers: SAR and the Convergence Sublayer (CS). The SAR reconstructs data that has been segmented into different cells (reassembly). It is also responsible for segmenting data that cannot fit within a 48-byte payload of an ATM cell (segmentation).

The CS determines the class of service to be used for a transmission. This will depend on the bit rate (constant or variable bit rate), the type of data, and the type of service to be provided (connection-oriented or connectionless). The quality of service parameters necessary for the transmission are determined by the class of service assigned.

6.5.3 High Speed LAN

6.5.3.1 Fast Ethernet and Gigabit Ethernet

It is nearly impossible to discuss networking without the mention of Ethernet, Fast Ethernet and Gigabit Ethernet. But, in order to determine which form is needed for your application, it’s important to first understand what each provides and how they work together.

A good starting point is to explain what Ethernet is. Simply, Ethernet is a very common method of networking computers in a LAN using copper cabling. Capable of providing fast and constant connections, Ethernet can handle about 10Mbps and can be used with almost any kind of computer.

While that may sound fast to those less familiar with networking, there is a very strong demand for even higher transmission speeds, which has been realized by the Fast Ethernet and Gigabit Ethernet specifications (IEEE 802.3u and IEEE 802.3z respectively). These LAN (local area network) standards have raised the Ethernet speed limit from 10 megabits per second (Mbps) to 100Mbps for Fast Ethernet and 1000Mbps for Gigabit Ethernet with only minimal changes made to the existing cable structure.

The building blocks of today's networks call out for a mixture of legacy 10BASE-T Ethernet networks and the new protocols. Typically, 10Mbps networks utilize Ethernet switches to improve the overall efficiency of the Ethernet network. Between Ethernet switches, Fast Ethernet repeaters are used to connect a group of switches together at the higher 100 Mbps rate.

However, with an increasing number of users running 100Mbps at the desktop, servers and aggregation points such as switch stacks may require even greater bandwidth. In this case, a Fast Ethernet backbone switch can be upgraded to a Gigabit Ethernet switch which supports multiple 100/1000 Mbps switches. High performance servers can be connected directly to the backbone once it has been upgraded.

6.5.3.2 Integrating Fast Ethernet and Gigabit Ethernet

Many client/server networks suffer from too many clients trying to access the same server, which creates a bottleneck where the server attaches to the LAN. Fast Ethernet, in combination with switched Ethernet, can create an optimal cost-effective solution for avoiding slow networks since most 10/100Mbps components cost about the same as 10Mbps-only devices.

When integrating 100BASE-T into a 10BASE-T network, the only change required from a wiring standpoint is that the corporate premise distributed wiring system must now include Category 5 (CAT5) rated twisted pair cable in the areas running 100BASE-T. Once rewiring is completed, gigabit speeds can also be deployed even more widely throughout the network using standard CAT5 cabling.

There is yet another variation of Ethernet called full-duplex Ethernet. Full-duplex Ethernet enables the connection speed to be doubled by simply adding another pair of wires and removing collision detection; the Fast Ethernet standard allowed full-duplex Ethernet. Until then all Ethernet worked in half-duplex mode which meant if there were only two stations on a segment, both could not transmit simultaneously. With full-duplex operation, this was now possible. In the terms of Fast Ethernet, essentially 200Mbps of throughput is the theoretical maximum per

full-duplex Fast Ethernet connection. This type of connection is limited to a node-to-node connection and is typically used to link two Ethernet switches together.

A Gigabit Ethernet network using the 1000BASE-LX long wavelength option supports duplex links of up to 550 meters of 62.5 millimeters or 50 millimeters multimode fiber. 1000BASE-LX can also support up to 5 Kilometers of 10 millimeter single-mode fiber. Its wavelengths range from 1270 millimeters to 1355 millimeters. The 1000BASE-SX is a short wavelength option that supports duplex links of up to 275 meters using 62.5 millimeters at multimode or up to 550 meters using 55 millimeters of multimode fiber. Typical wavelengths for this option are in the range of 770 to 860 nanometers.

The Fast Ethernet specification calls for two types of transmission schemes over various wire media. The first is 100BASE-TX, which, from a cabling perspective, is very similar to 10BASE-T. It uses CAT5-rated twisted pair copper cable to connect various hubs, switches and end-nodes. It also uses an RJ45 jack just like 10BASE-T and the wiring at the connector is identical. These similarities make 100BASE-TX easier to install and therefore the most popular form of the Fast Ethernet specification.

The second variation is 100Base-FX which is used primarily to connect hubs and switches together either between wiring closets or between buildings. 100BASE-FX uses multimode fiber-optic cable to transport Fast Ethernet traffic.

Gigabit Ethernet specification calls for three types of transmission schemes over various wire media. Gigabit Ethernet was originally designed as a switched technology and used fiber for uplinks and connections between buildings. Because of this, in June 1998 the IEEE approved the Gigabit Ethernet standard over fiber: 1000BASE-LX and 1000BASE-SX.

The next Gigabit Ethernet standardization to come was 1000BASE-T, which is Gigabit Ethernet over copper. This standard allows one gigabit per second (Gbps) speeds to be transmitted over CAT5 cable and has made Gigabit Ethernet migration easier and more cost-effective than ever before.

6.5.3.3 Basic Building Blocks for Fast Ethernet LAN

The basic building block for the Fast Ethernet LAN is the Fast Ethernet repeater. The two types of Fast Ethernet repeaters offered on the market today are:

Class I Repeater -- The Class 1 repeater operates by translating line signals on the incoming port to a digital signal. This allows the translation between different types of Fast Ethernet such as 100BASE-TX and 100BASE-FX. A Class I repeater introduces delays when performing this conversion such that only one repeater can be put in a single Fast Ethernet LAN segment.

Class II Repeater -- The Class II repeater immediately repeats the signal on an incoming port to all the ports on the repeater. Very little delay is introduced by this quick movement of data across the repeater; thus two Class II repeaters are allowed per Fast Ethernet segment.

Network managers understand the 100 meter distance limitation of 10BASE-T and 100BASE-T Ethernet and make allowances for working within these limitations. At the higher operating speeds, Fast Ethernet and 1000BASE-T are limited to 100 meters over CAT5-rated cable. The EIA/TIA cabling standard recommends using no more than 90 meters between the equipment in the wiring closet and the wall connector. This allows another 10 meters for patch cables between the wall and the desktop computer.

In contrast, a Fast Ethernet network using the 100BASE-FX standard is designed to allow LAN segments up to 412 meters in length. Even though fiber-optic cable can actually transmit data greater distances (i.e. 2 Kilometers in FDDI), the 412 meter limit for Fast Ethernet was created to allow for the round trip times of packet transmission. Typical 100BASE-FX cable specifications call for multimode fiber-optic cable with a 62.5 micron fiber-optic core and a 125 micron cladding around the outside. This is the most popular fiber optic cable type used by many of the LAN standards today. Connectors for 100BASE-FX Fast Ethernet are typically ST connectors (which look like Ethernet BNC connectors).

Many Fast Ethernet vendors are migrating to the newer SC connectors used for ATM over fiber. A rough implementation guideline to use when determining the maximum distances in a Fast Ethernet network is the equation: 400 - (r x 95)

where r is the number of repeaters. Network managers need to take into account the distance between the repeaters and the distance between each node from the repeater.

6.5.3.4 Maintaining a Quality Network

The CAT5 cable specification is rated up to 100 megahertz (MHz) and meets the requirement for high speed LAN technologies like Fast Ethernet and Gigabit Ethernet. The EIA/TIA (Electronics industry Association/Telecommunications Industry Association) formed this cable standard which describes performance the LAN manager can expect from a strand of twisted pair copper cable. Along with this specification, the committee formed the EIA/TIA-568 standard named the “Commercial Building Telecommunications Cabling Standard” to help network managers install a cabling system that would operate using common LAN types (like Fast Ethernet). The specification defines Near End Crosstalk (NEXT) and attenuation limits between connectors in a wall plate to the equipment in the closet. Cable analyzers can be used to ensure accordance with this specification and thus guarantee a functional Fast Ethernet or Gigabit Ethernet network.

The basic strategy of cabling Fast Ethernet systems is to minimize the re-transmission of packets caused by high bit-error rates. This ratio is calculated using NEXT, ambient noise and attenuation of the cable.

6.5.3.5 Fast Ethernet Migration

Most network managers have already migrated from 10BASE-T or other Ethernet 10Mbps variations to higher bandwidth networks. Fast Ethernet ports on Ethernet switches are used to provide even greater bandwidth between the workgroups at 100Mbps speeds. New backbone switches have been created to offer support for 1000Mbps Gigabit Ethernet uplinks to handle network traffic. Equipment like Fast Ethernet repeaters will be used in common areas to group Ethernet switches together with server farms into large 100Mbps pipes. This is currently the most cost effective method of growing networks within the average enterprise.

6.5.3.6 Fibre Channel

Fibre Channel is a gigabit-speed network technology primarily used for storage networking. It started for use primarily in the supercomputer field, but has become the standard connection type for storage area networks in enterprise storage. Despite its name, Fibre Channel signaling can run on both twisted-pair copper wire and fiber optic cables. Fibre Channel Protocol (FCP) is the interface protocol of SCSI (Small Computer System Interface) on the Fibre Channel.

6.5.3.7 Fibre Channel topologies

There are three major Fibre Channel topologies,

Point-to-Point (FC-P2P)- Two devices are connected back to back. This is the simplest topology, with limited connectivity.

Arbitrated loop (FC-AL)- In this design, all devices are in a loop or ring, similar to token ring networking. Adding or removing a device from the loop causes all activity on the loop to be interrupted. The failure of one device causes a break in the ring. Fibre Channel hubs exist to connect multiple devices together and may bypass failed ports. A loop may also be made by cabling each port to the next in a ring. Often an arbitrated loop between two ports will negotiate to become a P2P connection, but this is not required by the standard.

Switched fabric (FC-SW)- All devices or loops of devices are connected to Fibre Channel switches, similar conceptually to modern Ethernet implementations. The switches manage the state of the fabric, providing optimized interconnections. Very limited security is available in today's fibre channel switches.

Attribute	Point-to-Point	Arbitrated Loop	Switched Fabric
Max Ports	2	127	~16777216 (2^{24})
Max Bandwidth	2× Link Rate	2× Link Rate	(Number of Ports) × Link Rate
Address Size	N/A	8-bit ALPA	24-bit Port ID
Address Assignment	N_Port Login	Loop Init. and Fabric Login	Fabric Login
Concurrent Connections	1	1	Switch Ports/2
Effect of Port Failure	Link Fails	Loop Fails unless bypassed	Switch and Port Link Fails

Concurrent Maintenance	Link Down	May Disrupt Entire Loop	Switch and Port Link Down
Expansion	Additional P2P Links	Attach loop to Fabric	Expand Fabric
Redundancy	Add Redundant P2P Link	Use Dual Loops	Use Redundant Switches
Link Rates Supported	All	All (all devices must be same)	All (Mixed rates available)
Media Types Supported	All	All	All
Classes of Service Supported	All	1, 2, & 3	All
Frame Delivery	In Order	In Order	Not Guaranteed
Access to Medium	Dedicated	Arbitrated	Dedicated
Cost per Port	Port Cost	Port Cost + Loop Function	Port Cost + Fabric Port

Table 1 Fibre Channel Topologies

6.5.3.8 Fibre Channel layers

Fibre Channel is a layered protocol. It consists of 5 layers, namely:

FC0- The physical layer, which includes cables, fiber optics, connectors, pinouts etc.

FC1- The data link layer, which implements the 8b/10b encoding and decoding of signals.

FC2- The network layer, defined by the FC-PI-2 standard, consists of the core of Fibre Channel, and defines the main protocols.

FC3- The common services layer, a thin layer that could eventually implement functions like encryption or RAID.

FC4- The Protocol Mapping layer. Layer in which other protocols, such as SCSI, are encapsulated into an information unit for delivery to FC2.

FC0, FC1, and FC2 are also known as FC-PH, the physical layers of fibre channel.

Fibre Channel products are available at 1 Gbit/s, 2 Gbit/s and 4 Gbit/s. An 8 Gbit/s standard is being developed. A 10 Gbit/s standard has been ratified, but is currently only used to interconnect switches. No 10 Gbit/s initiator or target

products are available yet based on that standard. Products based on the 1, 2, 4 and 8 Gbit/s standards should be interoperable, and backward compatible; the 10 Gbit/s standard, however, will not be backward compatible with any of the slower speed devices.

6.5.3.9 Ports

The following ports are defined by Fibre Channel:

E_port is the connection between two fibre channel switches. Also known as an Expansion port. When E_ports between two switches form a link, that link is referred to as an InterSwitch Link or ISL.

EX_port is the connection between a fibre channel router and a fibre channel switch. On the side of the switch it looks like a normal E_port, but on the side of the router it is a EX_port.

F_port is a fabric connection in a switched fabric topology. Also known as Fabric port. An F_port is not loop capable.

FL_port is the fabric connection in a public loop for an arbitrated loop topology. Also known as Fabric Loop port. Note that a switch port may automatically become either an F_port or an FL_port depending on what is connected.

G_port or generic port on a switch can operate as an E_port or F_port.

L_port is the loose term used for any arbitrated loop port, NL_port or FL_port. Also known as Loop port.

N_port is the node connection pertaining to hosts or storage devices in a Point-to-Point or switched fabric topology. Also known as Node port.

NL_port is the node connection pertaining to hosts or storage devices in an arbitrated loop topology. Also known as Node Loop port.

TE_port is a term used for multiple E_ports trunked together to create high bandwidth between switches. Also known as Trunking Expansion port.

6.5.3.10 Fibre Channel Infrastructure

Fibre Channel switches are divided into two classes of switches. These classes are not part of the standard, and the classification of every switch is left up to the manufacturer.

Director switches are characterized by offering a high port-count in a modular (slot-based) chassis with no single point of failure (high availability).

Fabric switches are typically fixed-configuration (sometimes semi-modular) non-redundant switches.

Brocade, Cisco and McData provide both Director and fabric switches. QLogic provides fabric switches. If multiple switch vendors are used in the same fabric, the fabric will default to “interoperability mode” where some proprietary advanced features may be disabled.

6.5.3.11 High Speed Wireless LAN

The Wireless Local Area Network (WLAN) industry has emerged as one of the fastest-growing segments of the communications industry. This growth was due, in large part, to the introduction of standards-based WLAN products. These products – based on the 802.11b standard – are faster, lower in cost, and simpler to setup and use than previous generation products. The majority of WLAN products today communicate at speeds up to 11 megabits per second (Mbps).

Two new WLAN standards are now emerging and will deliver higher speeds, up to 54 Mbps, to WLAN users. These new standards are known as 802.11a and 802.11g.

6.5.3.12 802.11a – High-Speed, High Capacity

By moving to the 5 GHz frequency band and by using OFDM (Orthogonal Frequency-Division Multiplexing) modulation, the 802.11a standard provides two key benefits over 802.11b. It increases the maximum speed per channel (from 11 Mbps to 54 Mbps) and increases the number of non-overlapping channels. The 5 GHz band (also known as the UNII band) is actually made up of three sub-bands, UNII1 (5.15-5.25 GHz), UNII2 (5.25-5.35 GHz) and UNII3 (5.725-5.825 GHz). Up to 8 non-overlapping channels are available when UNII1 and UNII2 are both used, versus 3 in the 2.4 GHz band. The total bandwidth available in the 5 GHz band is also higher than in the 2.4 GHz band – 83.5 MHz versus 300 MHz. Thus, an 802.11a-based WLAN can support a larger number of simultaneous high-speed users without the potential for conflict. These benefits come, however, with some tradeoffs in terms of compatibility and range. Because they operate in

different frequency bands, 802.11a and 802.11b products are not compatible. A 2.4 GHz 802.11b access point, for example, won't work with a 5 GHz 802.11a network card. However, both standards can certainly co-exist. For example, an 802.11a user and an 802.11b user, using separate access points and clients for each, connected to the same LAN, can operate in the same physical space and share network resources including broadband and internet access.

The higher operating frequency of 802.11a equates to a relatively shorter range. You will need a larger number of 802.11a access points to cover the same area. The FCC (Federal Communications Commission) requires a max 6dB antenna attached to the radio when UNII1 and UNII2 are both used to get a full 8 indoor channels, reducing range. However, initial tests show that 802.11a products still maintain about a 3 to 1 performance improvement versus 802.11b over typical indoor ranges.

6.5.3.13 802.11g – High-Speed in the 2.4 GHz Band

The 802.11g standard brings the benefits of higher speeds, while maintaining backward compatibility with existing 802.11b equipment. 802.11g specifies operation in the same 2.4 GHz frequency band and with the same DSSS modulation types as 802.11b at speeds up to 11 Mbps, while adding more efficient OFDM modulation types at higher speeds. An 802.11g network card, for example, will work with an 802.11b access point and an 802.11g access point will work with 802.11b network cards – at speeds up to 11 Mbps. To benefit from higher speeds up to 54 Mbps, both the access point and network card must be 802.11g compliant. The draft standard also specifies optional modulation types (OFDM/CCK) that are intended to improve efficiency in an all-802.11g installation.

The tradeoff with 802.11g is in a lower capacity, versus 802.11a, to serve a large number of high speed WLAN users. The OFDM modulations allow for higher speed but the total available bandwidth in the 2.4 GHz frequency band remains the same because 802.11g is still restricted to three channels in the 2.4 GHz band, unlike the eight that are available in the 5Ghz band.

Choosing a High-Speed WLAN Standard

Each WLAN deployment is unique. It is impossible to provide a simple answer to choosing between 802.11a or 802.11g. In some cases, it may even make sense to mix both, especially as the industry introduces dual-mode (two radio) solutions that support all three standards. In the next few years, the industry is even expected to ship multi-mode devices that support 11a, 11b and 11g.

It is important to evaluate both your anticipated needs and your current installation. Some key criteria for choosing a high-speed WLAN standard are:

Total Capacity Requirements: For installations that call for high-density populations of mostly high-speed WLAN users, 802.11a may be a better choice. If a smaller number of high-speed users are being added to an existing 802.11b installation, 802.11g may be the better choice.

Timing of High-Speed Need: If high-speed is needed immediately, 802.11a would be the way to go, since 802.11g products are not expected to be shipped until the second half of 2002.

Migration Plan for Existing Installation: If you have a large 802.11b installation and simply want add a few high-speed users in the next year, it may make sense to try both options now and plan to deploy dual-mode products in the following year.

Interference Concerns: If you are currently experience interference in the 2.4 GHz frequency band from products like Bluetooth and cordless phones, it may make sense to move to the less crowded (for now) 5 GHz band with 802.11a.

6.6 Summary

A frame relay network gains efficiency by trading network complexity off against link capacity. Within a frame relay network link and node utilizations have to be kept low in order for the network to operate stably. The network and end systems must provide a low –latency, high bandwidth path between applications to support low inter-application delay.

ATM allows multiple services on single network and uniform technology. Fibre channel protocol can run on both twisted pair copper and fiber optic cables.

The 802.11a/b/g standards are commonly used in Laptops for wireless networking.

6.7 Keywords

SNA- Systems Network Architecture (SNA) is IBM's proprietary networking architecture created in 1974. It is a complete protocol stack for interconnecting computers and their resources. SNA describes the protocol and is, in itself, not actually a program. The implementation of SNA takes the form of various communications packages, most notably VTAM which is the mainframe package for SNA communications. SNA is still used extensively in banks and other financial transaction networks, as well as in many government agencies.

DECnet- It is a proprietary suite of network protocols created by Digital Equipment Corporation, originally released in 1975 in order to connect two PDP-11 minicomputers. It evolved into one of the first peer-to-peer network architectures, thus making DEC into a networking powerhouse in the 1980s.

FDDI- Fiber-Distributed Data Interface (FDDI) provides a standard for data transmission in a local area network that can extend in range up to 200 kilometers (124 miles). The FDDI protocol uses as its basis the token ring protocol. In addition to covering large geographical areas, FDDI local area networks can support thousands of users.

X.25- X.25 is an ITU-T (International Telecommunication Union - Telecommunication) standard protocol suite for wide area networks using the phone or ISDN system as the networking hardware.

T1- Digital signal 1 (DS1, also known as T1, sometimes "DS-1") is a T-carrier signaling scheme devised by Bell Labs. DS1 is a widely used standard in telecommunications in North America and Japan to transmit voice and data between devices.

SONET-Synchronous optical networking is a method for communicating digital information using lasers or light-emitting diodes (LEDs) over optical fiber.

100BASE-T -100BASE-T is any of several Fast Ethernet 100 Mbit/s (12.5 MByte/s excluding 4B/5B overhead) CSMA/CD standards for twisted pair cables,

including: 100BASE-TX (100 Mbit/s over two-pair Cat5 or better cable), 100BASE-T4 (100 Mbit/s over four-pair Cat3 or better cable, defunct), 100BASE-T2 (100 Mbit/s over two-pair Cat3 or better cable, also defunct). The segment length for a 100BASE-T cable is limited to 100 meters (as with 10BASE-T and gigabit Ethernet). All are or were standards under IEEE 802.3

Optical Carrier levels (OCx)- The Synchronous Optical Network (SONET) includes a set of signal rate multiples for transmitting digital signals on optical fiber. The base rate (OC-1) is 51.84 Mbps. Certain multiples of the base rate are provided as shown in the following table. Asynchronous transfer mode (ATM) makes use of some of the Optical Carrier levels.

OC-1 means Data rate is 51.84 Mbps, OC-3 means 155.52 Mbps and OC-12 means 622.08 Mbps.

6.8 Self Assessment Questions

Q1. Explain the need of high speed networking.

Q2. Explain different performance attributes considered in high speed networks.

Q3. Explain the architecture of ATM networks.

Q4. Explain fibre channel layers.

Q5. What are the basic building blocks of fast Ethernet.

Q6. Explain frame relay devices and its network implementation.

Q7. Explain different standards in high speed wireless LAN.

6.9 References

1. Wide Area High Speed Networking, Sidni Feit, Macmillan Technical Publishing

2. High Speed Networks and Internets Performance and Quality of Service, Second Edition, William Stalling, Pearson Education

3. High Speed Networking, James P.G. Sterbenz, Joseph D. Touch, Willey Computer Publishing

4. International Technical Support Organization High-Speed Networking Technology: An Introductory Survey, IBM International Technical Support Centre

Paper Code: MCA- 405	**Author: Parvinder Singh**
Paper Name: Computer Networks-II	**Vetter: Dinesh Kumar**
Lesson Number: 07	

Applications of Multimedia

Structure

7.0 Objective

As the name of the chapter implies that objective of this chapter is to deal with real life applications of multimedia. In this chapter various user activities and multimedia files format are also described. Finally one of the very interesting applications of multimedia – video conferencing is explained.

7.1 User Activities and Intra- activities

Multimedia is everything you can hear or see: texts, books, pictures, music, sounds, CDs, videos, DVDs, Records, Films, animations and more. To

understand the user activities it is better to know the formats of multimedia available for users.

7.1.1 Multimedia Formats

Multimedia comes in many different formats. Multimedia elements (like sounds or videos) are stored in media files. The most common way to discover the media type is to look at the file extension. When a browser sees the file extensions .htm or .html, it will assume that the file is an HTML page. The .xml extension indicates an XML file, and the .css extension indicates a style sheet. Picture formats are recognized by extensions like .gif and .jpg. Multimedia elements also have their own file formats with different extensions.

Sound can be stored in many different formats.

7.1.1.1 MIDI Format

The MIDI (Musical Instrument Digital Interface) is a format for sending music information between electronic music devices like synthesizers and PC sound cards.

The MIDI format was developed in 1982 by the music industry. The MIDI format is very flexible and can be used for everything from very simple to real professional music making.

MIDI files do not contain sampled sound, but a set of digital musical instructions (musical notes) that can be interpreted by your PC's sound card.

The downside of MIDI is that it cannot record sounds (only notes). Or, to put it another way: It cannot store songs, only tunes.

The upside of the MIDI format is that since it contains only instructions (notes), MIDI files can be extremely small. The example above is only 23K in size but it plays for nearly 5 minutes.

The MIDI format is supported by many different software systems over a large range of platforms. MIDI files are supported by all the most popular Internet browsers. Sounds stored in the MIDI format have the extension .mid or .midi.

7.1.1.2 RealAudio Format

The RealAudio format was developed for the Internet by Real Media. The format also supports video. The format allows streaming of audio (on-line music, Internet radio) with low bandwidths. Because of the low bandwidth priority, quality is often reduced. Sounds stored in the RealAudio format have the extension .rm or .ram. However, the latest version of RealProducer, Real's flagship encoder, reverted to using .ra for audio-only files, and began using .rv for video files (with or without audio), and .rmvb for VBR (Variable bitrate) video files.

The .ram (Real Audio Metadata) and .smil (Synchronized Multimedia Integration Language) file formats are sometimes encountered as links from web pages

7.1.1.3 AU Format

The AU format is supported by many different software systems over a large range of platforms. Sounds stored in the AU format have the extension .au. The format was introduced by Sun Microsystems. It consists of a header of 5 32-bit words, an optional information chunk and then the data (in big endian format). In big endian, most significant byte (MSB) value is stored at the memory location with the lowest address, the next byte value in significance, is stored at the following memory location and so on. This is akin to Left-to-Right reading order in hexadecimal.

7.1.1.4 AIFF Format

The AIFF (Audio Interchange File Format) was developed by Apple. AIFF files are not cross-platform and the format is not supported by all web browsers. Sounds stored in the AIFF format have the extension .aif or .aiff. AIFF is also used by Silicon Graphics Incorporated.

The audio data in a standard AIFF file are uncompressed big-endian pulse-code modulation (PCM). There is also a compressed variant of AIFF known as AIFF-C or AIFC, with various defined compression codecs.

Standard AIFF is a leading format (along with SDII and WAV) used by professional-level audio and video applications, as unlike the better known consumer MP3 format, it is non-compressed (which aids rapid streaming of multiple audio files from disk to the application), and lossless. Like any non-compressed, lossless format, it uses much more disk space than MP3 -- about 10MB for one minute of stereo audio. In addition to audio data, AIFF can include loop point data and the musical note of a sample, for use by hardware samplers and musical applications.

Traditional AIFF-C compressed formats are of poor quality and were used only when necessary in multimedia applications. With the development and popularization of the much higher quality MP3, and related compressed audio formats, their use has become essentially nonexistent.

7.1.1.5 SND Format

The SND (Sound) was developed by Apple. SND files are not cross-platform and the format is not supported by all web browsers. Sounds stored in the SND format have the extension .snd.

7.1.1.6 WAVE Format

The WAVE (waveform) format is developed by IBM and Microsoft. It is supported by all computers running Windows, and by all the most popular web browsers.
Sounds stored in the WAVE format have the extension .wav.
Though a WAV file can hold compressed audio, the most common WAV format contains uncompressed audio in the pulse-code modulation (PCM) format. PCM audio is the standard audio file format for CDs at 44,100 samples per second, 16 bits per sample. Since PCM uses an uncompressed, lossless storage method, which keeps all the samples of an audio track, professional users or audio experts may use the WAV format for maximum audio quality. WAV audio can also be edited and manipulated with relative ease using software.

Uncompressed WAV files are quite large in size, so, as file sharing over the Internet has become popular, alongside miniature portable music players such as the iPod, the WAV format has declined in popularity as a format for transmission. However, it is still a commonly used, relatively "pure", i.e. lossless, file type, suitable for retaining "first generation" archived files of high quality, or use on a system where high fidelity sound is required and disk space is not restricted.

7.1.1.7 MP3 Format (MPEG)

MP3 files are actually MPEG files. But the MPEG format was originally developed for video by the Moving Pictures Experts Group. We can say that MP3 files are the sound part of the MPEG video format.

MP3 is one of the most popular sound formats for music recording. The MP3 encoding system combines good compression (small files) with high quality. Expect all your future software systems to support it.

Sounds stored in the MP3 format have the extension .mp3, or .mpga (for MPG Audio).

It is a lossy compression format, and designed to greatly reduce the amount of data required to represent audio, yet still sound like a faithful reproduction of the original uncompressed audio to most listeners. It was invented by a team of European engineers of Philips, CCETT (Centre commun d'études de télévision et télécommunications), IRT and Fraunhofer Society, who worked in the framework of the EUREKA 147 DAB digital radio research program, and it became an ISO/IEC (International Organization for Standardization/International Electrotechnical Commission) standard in 1991.

7.1.1.8 Sound Format on Internet

The WAVE format is one of the most popular sound formats on the Internet, and it is supported by all popular browsers. If you want recorded sound (music or speech) to be available to all your visitors on your web page, you should use the WAVE format.

The MP3 format is the new and upcoming format for recorded music. If your website is about recorded music, the MP3 format is the choice of the future.
Video can be stored in many different formats.

7.1.1.9 AVI Format

The AVI (Audio Video Interleave) format was developed by Microsoft. The AVI format is supported by all computers running Windows, and by all the most popular web browsers. It is a very common format on the Internet, but not always possible to play on non-Windows computers. Videos stored in the AVI format have the extension .avi.
AVI files can contain both audio and video data in a standard container that allows synchronous audio-with-video playback. Like DVDs, AVI files support multiple streaming audio and video, although these features are seldom used.

7.1.1.10 Windows Media Format

The Windows Media format is developed by Microsoft. Windows Media is a common format on the Internet, but Windows Media movies cannot be played on non-Windows computer without an extra (free) component installed. Some later Windows Media movies cannot play at all on non-Windows computers because no player is available. Videos stored in the Windows Media format have the extension .wmv.

7.1.1.11 MPEG Format

The MPEG (Moving Pictures Expert Group) format is the most popular format on the Internet. It is cross-platform, and supported by all the most popular web browsers. Videos stored in the MPEG format have the extension .mpg or .mpeg.
MPEG is a working group of ISO/IEC charged with the development of video and audio encoding standards.
MPEG has standardized the following compression formats and ancillary standards:

MPEG-1: Initial video and audio compression standard. Later used as the standard for Video CD, and includes the popular Layer 3 (MP3) audio compression format.

MPEG-2: Transport, video and audio standards for broadcast-quality television. Used for over-the-air digital television ATSC (Advanced Television Systems Committee), DVB (Digital Video Broadcasting) and ISDB (Integrated Services Digital Broadcasting), digital satellite TV services like Dish Network, digital cable television signals, and with slight modifications for DVDs.

MPEG-3: Originally designed for HDTV, but abandoned when it was discovered that MPEG-2 (with extensions) was sufficient for HDTV. (Do not confuse with MP3, which is MPEG-1 Layer 3.)

MPEG-4: Expands MPEG-1 to support video/audio "objects", 3D content, low bitrate encoding and support for Digital Rights Management. Several new (newer than MPEG-2 Video) higher efficiency video standards are included (an alternative to MPEG-2 Video), notably: MPEG-4 Part 2 (or Advanced Simple Profile) and MPEG-4 Part 10 (or Advanced Video Coding or H.264). MPEG-4 Part 10 may be used on HD DVD and Blue-ray discs, along with VC-1 and MPEG-2.

In addition, the following standards, while not sequential advances to the video encoding standard as with MPEG-1 through MPEG-4, are referred to by similar notation:

MPEG-7: A formal system for describing multimedia content.

MPEG-21: MPEG describes this standard as a multimedia framework.

7.1.1.12 QuickTime Format

The QuickTime format is developed by Apple. QuickTime is a common format on the Internet, but QuickTime movies cannot be played on a Windows computer without an extra (free) component installed. Videos stored in the QuickTime format have the extension .mov.

The QuickTime (.mov) file format functions as a multimedia container file that contains one or more tracks, each of which stores a particular type of data:

audio, video, effects, or text (for subtitles, for example). Each track either contains a digitally-encoded media stream (using a specific codec) or a data reference to the media stream located in another file. Tracks are maintained in a hierarchal data structure consisting of objects called atoms. An atom can be a parent to other atoms or it can contain media or edit data, but it cannot do both. The ability to contain abstract data references for the media data, and the separation of the media data from the media offsets and the track edit lists means that QuickTime is particularly suited for editing, as it is capable of importing and editing in place (without data copying).

7.1.1.13 Real Video Format

The Real Video format was developed for the Internet by Real Media. The format allows streaming of video (on-line video, Internet TV) with low bandwidths. Because of the low bandwidth priority, quality is often reduced. Videos stored in the Real Video format have the extension .rm or .ram.

7.1.1.14 Shockwave (Flash) Format

The Shockwave format was developed by Macromedia. The Shockwave format requires an extra component to play. This component comes preinstalled with the latest versions of Netscape and Internet Explorer. Videos stored in the Shockwave format have the extension .swf.

Windows media files have the extensions: .asf, .asx, .wma, and .wmv.

7.1.1.15 ASF Format

The ASF format (Advanced Streaming Format) is specially designed to run over the Internet. ASF files can contain audio, video, slide shows, and synchronized events. ASF files can be highly compressed and can be delivered as a continuous flow of data (on-line TV or radio). Files can be of any size, and can be compressed to match many different bandwidths (connection speeds).

7.1.1.16 ASX Format

ASX (Advanced Stream Redirector) files are not media files, but metafiles.

Metafiles provides information about files. ASX files are plain text files used to describe multimedia content.

The Advanced Stream Redirector (ASX) format is a type of XML metafile designed to store a list of Windows Media files to play during a multimedia presentation.

It is used frequently on streaming video servers where multiple ASF files are to be played in succession.

7.1.1.17 WMA Format

The WMA (Windows Media Audio) format is an audio format developed by Microsoft. WMA is designed to handle all types of audio content. The files can be highly compressed and can be delivered as a continuous flow of data (on-line radio). WMA files can be of any size, and be compressed to match many different bandwidths (connection speeds). The WMA format is similar to the ASF format.

7.1.1.18 WMV Format

The WMV (Windows Media Video) format is a video format developed by Microsoft. WMV is designed to handle all types of video content. The files can be highly compressed and can be delivered as a continuous flow of data (on-line radio). WMV files can be of any size, and be compressed to match many different bandwidths (connection speeds). The WMV format is similar to the ASF format.

WMV (*.wmv) files use Microsoft's Advanced Systems Format (ASF) container format (digital). These files can be played by players such as MPlayer or Windows Media Player, the latter being only available for Microsoft Windows and Macintosh systems. Many third-party players exist for various platforms such as Linux that use the FFmpeg implementation of the WMV codecs.

The extension .wmv typically describes ASF files that use Windows Media Video codecs. The audio codec used in conjunction with WM Video is typically some form of Windows Media Audio, or in rarer cases the somewhat deprecated Sipro ACELP.net audio codec. Microsoft recommends that ASF files containing non-Windows Media codecs use the more generic .asf file extension.

7.1.1.19 Other Windows Media Formats

WAX (Windows Media Audio Redirector) files are much the same as ASX files, but intended to describe audio files (.wma files). WMP (Windows Media Player) files and WMX are reserved file types for future use by Windows.

7.1.2 User's activities

From user's point of view, multimedia may be broadly divided into linear and non-linear categories. Linear active content progresses without any navigation control for the viewer such as a cinema presentation.

Non-linear content offers user interactivity to control progress as used with a computer game or used in self-paced computer based training. Non-linear content is also known as hypermedia content. In hypermedia content user can visit to other information by clicking on the link provided by means of text, message, image or animation.

7.2 Stand Alone Multimedia Applications

The various stand alone applications of multimedia are described below.

7.2.1 Advertising

The advertising in one of the very popular fields where multimedia can be used. As we see in different advertisements, visual and sound effects have long lasting impression on viewers, so multimedia plays very important role in this field.

7.2.2 Business Presentations

In business presentations, multimedia is used to display images, shapes, tables, text, graphs and animations. By including various multimedia components, one

can make very effective presentation. Another advantage of using multimedia is that, there is no need to explain the things by lengthy speech or text, as usage of various multimedia components make the content self explanatory.

7.2.3 Computer Simulations

In computer simulation there is no need of actual hardware, components and devices as various images and shapes can be made to behave like them by using multimedia and programming. One can easily and quickly perform various experiments using this way.

7.2.4 Teaching

Teachers make the slides of their lectures to deliver more effectively the contents. It is also possible to demonstrate and show the visuals, which was otherwise not possible.

7.2.5 Entertainment

The animations and movies are the examples of multimedia usage in entertainment. The dangerous and hart blowing effects which are included in movies are possible due to multimedia. video games are also examples of multimedia in entertainment.

7.2.6 Virtual Reality

In virtual reality, one can see and feel like he is actually visiting that place or building. This is done by capturing lots of images and video and making hot spots in them so that user can follow in that direction or enter from that way. It is also used in games like racing where user feels he is actually driving the bike or cars.

Apart of it multimedia is also used for-

- Governmental Services
- Journalism
- Nonprofit Services
- Professional Training

- Software Interfaces
- Spatial Temporal Applications

7.3 Single User Networked Applications

Single user network applications are those applications where a user can access the multimedia database available on network, but he can't interact with other users on that multimedia applications.

7.3.1 Engineering

In Engineering, especially in mechanical and automobile engineering, multimedia is primarily used for designing a machine or an automobile. This lets an Engineer view a product from various perspectives, zoom in on critical parts and do other manipulations, before actually producing it. This is known as computer-aided design (CAD).

7.3.2 Medicine

In Medicine, doctors can get trained by looking at a virtual surgery or they can simulate how the human body is affected by diseases spread by viruses and bacteria and then develop techniques to prevent it.

7.3.3 Mathematical and Scientific Research

In Mathematical and Scientific Research, multimedia are mainly used for modeling and simulation. For example, a scientist can look at a molecular model of a particular substance and manipulate it to arrive at a new substance.

7.3.4 Arts

In the Arts there are multimedia artists, whose minds are able to blend techniques using different media that in some way incorporates interaction with the viewer.

7.3.5 Education

In Education, multimedia is used to produce computer-based training courses (popularly called CBTs) and reference books like encyclopedia and almanacs. A CBT lets the user go through a series of presentations, text about a particular topic, and associated illustrations in various information formats. Edutainment is an informal term used to describe combining education with entertainment, especially multimedia entertainment.

7.3.6 Industry

In the Industrial sector, multimedia is used as a way to help present information to bosses and co-workers.

7.3.7 Multimedia Messaging System

The Multimedia Messaging System, or MMS, is an application that allows one to send and receive messages containing Multimedia - related content. MMS is a common feature of most cell phones. An electronic multimedia encyclopedia can present information in better ways than traditional encyclopedia, so the user has more fun and learns more quickly. For instance, an article on World War II can include hyperlinks to articles on countries involved in the war. When users click on a hyperlink, they are redirected to a detailed article about that country. In addition, it can include a video on the related contents. It can also present maps pertinent to World War II. Hyperlinks let a user access information in a non-linear fashion as opposed to print materials which are essentially linear. This can speed-up learning and improve the user experience, when added to multiple elements such as pictures, photographs, audio and video. (It is also said that some people learn better by seeing than reading and some others by listening).

7.4 Audio Conferencing

As the telephone calls are made between two people, Internet service provides a group of users to audio discussion similar to a conference call. Each discussion is known as an audio conference. To create a conference, a user runs software

that organizes and controls the discussion. The software asks the name of people interested in conference, and then try to contact all of them one by one. The software do this by sending messages to them about the conference detail.

To join a conference, a user must run a program that handles audio reception and transmission. The program monitors the user's microphone, converts the signal to digital form, and sends a copy to other users in the conference. The program also receives messages, converts the messages back into sound, and plays the results for the user to hear. All participants hear the conversation similar to a conference telephone call.

Audio conferencing is significant because it allows participants to convey and understand emotion. Unlike written communication, voice carries inflection that tells the listener whether the speaker is excited, tired, angry, or joking.

7.5 Video Conferencing

Audio conferencing and other shared services provide a way for people to work together. Such services can be especially helpful when preparing or reviewing a document. However, face-to-face interaction usually works better when a group needs to generate ideas or discuss a topic because facial expressions can show surprise or agreement without requiring words.

To enable face-to-face interaction the Internet offers video conferencing services. A video conference begins as a session, when a user runs a program that starts a video session. The software allows the user to enter information about other participants, and contacts each of them. When a new participant joins the conference an image from the camera on their computer appears in a small window on everyone's screen. The pictures are similar to an ordinary television picture. When a participant smiles, everyone see his expression.

Of course, video alone does not help people communicate. Therefore, most video conferencing services incorporate both video and audio into a single conference session. When a participant joins the conference, they hear as well as see all other participants.

A video conference works well when used with a few people. Each user's display shows the pictures of other participants. However, when more than a few pictures appear on a display, it becomes difficult to watch all of them.

When many people participate in a video conference, the screen cannot hold individual images from cameras on the participants' computers. Instead, people must gather in smaller groups in rooms that each have a camera and a large-screen computer display. In essence, the entire room becomes a single participant in the conference. The camera sends a picture of the room to all other participants where it is projected on a display large enough for everyone to see. Instead of small rectangles that each contains an individual's face, the display contains rectangles that each shows a view of a remote room.

Sending audio for an entire room can be difficult. A single microphone may not be sensitive enough to detect sounds in all parts of the room. Multiple microphones sometimes detect background noise such as papers rustling. To handle the problem, rooms equipped for a video conference sometimes have portable microphones that can be passed to whoever needs to speak.

7.6 Summary

Multimedia on internet and network is one of the future applications where there are lots of possibilities to expand in various directions. Only thing is one have to think in some naval way. As we are seeing many commercial radio and TV stations have also begin to transmit 24 hrs online transmissions on internet. Only limitations which are there on computers - large size of multimedia files and low speed of internet.

7.7 Keywords

DVD- DVD (commonly "Digital Versatile Disc" or "Digital Video Disc") is an optical disc storage media format that can be used for data storage, including movies with high video and sound quality. DVDs resemble compact discs as their diameter is the same (120 mm (4.72 inches) or occasionally 80 mm (3.15 inches)

in diameter), but they are encoded in a different format and at a much higher density.

HTML- In computing, HyperText Markup Language (HTML) is the predominant markup language for the creation of web pages. It provides a means to describe the structure of text-based information in a document — by denoting certain text as headings, paragraphs, lists, and so on — and to supplement that text with interactive forms, embedded images, and other objects.

Hyperlink- A hyperlink (often referred to as simply a link), is a reference or navigation element in a document to another section of the same document, another document, or a specified section of another document, that automatically brings the referred information to the user when the navigation element is selected by the user. As such it is similar to a citation in literature, but with the distinction of automatic instant access.

Hypermedia- Hypermedia is a term created by Ted Nelson. It is used as a logical extension of the term hypertext, in which graphics, audio, video, plain text and hyperlinks intertwine to create a generally non-linear medium of information. This contrasts with the broader term multimedia, which may be used to describe non-interactive linear presentations as well as hypermedia.

CBT- Computer-based training (CBT), also called computer-assisted instruction (CAI) is a type of education in which the student learns by executing special training programs on a computer. CBT is especially effective for training people to use computer applications because the CBT program can be integrated with the applications so that students can practice using the application as they learn.

VBR- VBR files vary the amount of output data per time segment. VBR allows a higher bitrate (and therefore more storage space) to be allocated to the more complex segments of media files while less space is allocated to less complex segments. The average of these rates is calculated to produce an average bitrate for the file that will represent its overall sound quality. MP3, WMA, Vorbis, and AAC audio files can optionally be encoded in VBR. Variable bit rate encoding is also commonly used on MPEG-2 video.

7.8 Self Assessment Questions

Q1. Explain the various multimedia file formats.

Q2. Explain the stand alone multimedia applications.

Q3. How are single user networked applications different from multi-user network applications and stand alone multimedia applications?

Q4. Explain various single user networked multimedia applications.

Q5. Explain audio and video conference in detail.

Q6. Write short notes on –

i) Virtual Reality ii) MMS iii) CBT iv) Sound Format on Internet

7.9 References

1. Internet Books, Third Edition, Douglas E. Commer, Pearson Education.
2. Design for Multimedia Learning, Boyle T., Prentice Hall
3. Multimedia Projects in Education: Designing, Producing and Assessing, Karen S. Ivers, Ann E. Barron, Barnes and Nobles.

Paper Code: MCA- 405	**Author: Parvinder Singh**
Paper Name: Computer Networks-II	**Vetter: Dinesh Kumar**
Lesson Number: 08	

Multimedia Networking

Structure

8.0 Objective

The objective of this lesson is to discuss applications of multimedia networking. Apart of it this lesson also explains streaming in detail. The description of standard internet protocol, RTP for transportation of audio and video is also given.

8.1 Introduction

Networking of computer based multimedia is now possible. However network delivery of multimedia is not new, the broadcast television system has been available since the 1950's. The difference now is that the information which has previously been presented on radio, TV, or in books, is now being presented or can be accessed by computer networks. Or looking at this development in another way computer technology has almost caught up with the broadcasters. The drive to deliver multimedia information over a computer network is fuelled by Information Superhighway which needs multimedia to justify the requirement for the highway. However the multimedia makes heavy demands on storage and transmission systems. Data compression can be used to reduce the demands of multimedia, particularly of video and audio on these systems, but usually at the expense of some loss in the detail compared with the source and at extra cost.

The ways in which users or participants in multimedia sessions access multimedia or connect with others have important consequences for the storage and transmission systems. For instance multimedia learning material can be accessed directly from a server during a class or downloaded to student machines prior to a session. The demands on a connecting network are very

different in each access mode. The cost of transmitting multimedia information will determine the pace of development of networked multimedia applications. The availability of standards for multimedia networking, particularly for inter-working between applications, the development of networked applications, and inter-working between networks are essential to reduce the complexity and level of skill required in using multimedia.

8.2 Multimedia Networking Applications

A number of real time applications are available for use over networks based on computers, which are capable using audio or video. The two networks over which applications have been developed till date include ISDN, Local Area Networks, and Local Area Networks connected over high speed ATM or SMDS connections. The table below summarizes the most important applications with examples.

Application	**ISDN**	**LAN**
Desk Top Conferencing	Proprietary Desktop conferencing systems	wb... an Internet white board tool, proprietary software
Video Conferencing	H.320 video conferencing	vat ... AVisual audio internet tool, ivs... video conferencing software
Video Mail	Based on H.320 video conferencing	X.400 and MIME electronic mail with video content
Remote Image viewing and manipulation	Medical applications eg. X-Rays	Medical Applications, Multicast of JPEG Satellite images
Information Kiosks	Proprietary systems	
Distance Learning	File Server Access	File Server Access

Table 8.1 Applications of Multimedia in ISDN and LAN

8.2.1 Desk Top Conferencing

Desk Top Conferencing is a means of working with a remote user on their computer running the same application. Typically a drawing or spreadsheet will be transmitted via the ISDN line. A voice conversation between the two users can be held at the same time. Within desk top conferencing remote control of a computer is possible. Since these applications must reproduce the screen of a PC on another, the speed of the connection between the two computers is important. Products are available from Fujitsu for LAN and ISDN use and IBM (Person to Person). Desktop conferencing is also available over some of the video conferencing products such as the Olivetti PCC.

The standardization of multimedia conferencing using the T.120 series of International Telecommunication Union standards is now widely accepted. Products based on these standards are likely to appear on the market next year from several manufacturers. The applicability of the T.120 standards will not be limited to WAN or ISDN networks, but will be appropriate to Local Area Networks too. The T.120 standards will apply to terminal with audio, audio and interactive video, or interactive graphics or all three. It will support point to point and multi-point conferencing. ISDN is the initial focus. Work is underway to define ATM support. Areas to be standardized include still image, annotation, application sharing, conference control, and multi-point conferencing.

8.2.2 Video Conferencing or Videophone products

Until recently the only implementations of video over the telephone network have been poor quality video phones or very expensive video conferencing for executives. A selection of video conferencing products is listed below:

VidiMac for the Apple Mac based on the Planet ISDN card, which uses motion JPEG techniques. IBM, Fujitsu, and Olivetti have video conferencing products based on the BT VC8000 PC card.

The IBM product is called Screen Call, the Fujitsu is called Team Vision, and the Olivetti is called PCC. All three use the services of the VC8000 card in different

ways. The common feature is support for H.320 video conferencing. Support is available in all three for file transfer, whiteboards, and remote control. Intel have released a product called ProShare, which is very competitive, but at present will not work to the H.320 standard.

Northern Telecom have a product called Visit 2.0 which runs on a PC or Mac. It uses external ISDN Terminal adapters and audio transmission need to be via an independently set up telephone call.

Invision have made proprietary non H.320 video conferencing available for LAN connections running LAN protocols such as TCP/IP. Frame rates range from 1 to 20 frames per second with corresponding data rates fro 64 to 512 kbps.

Picturetel Live PCS100 is one of the more expensive PC based products, but manages excellent quality through good implementations and extensions to the H.320 standard. Full CIF pictures are available. At QCIF resolution of 7.5 frames per second are available.

The Olivetti PCC based on the BT VC8000 PC card provides QCIF at 15 frames per second, audio, file transfer, whiteboard, remote application control, remote form entry, image capture and transfer, and a text chat mode. All of these applications are programmable and should users should use this feature to customize the screen interface, which is based on standard Windows menu bars and buttons. Interfaces are available for video and audio from other sources. All of these facilities are programmable and customizable. H.320 standards are supported. Most of these products will provide ISDN applications such as file transfer, but the additional facility of a quarter screen video picture of the caller will be available in colour. Compression techniques are used to improve the quality of moving pictures. Sadly these some products use different compression methods so they cannot communicate with each other. But with manufacturers moving to the H.320 series of compression standards, at least for video phone communications this situation may improve. Other standards e.g. MPEG are be better for straight video broadcast.

8.2.3 Video Mail

Video mail can be delivered via X.400 mail, or Internet mail with MIME (Multipurpose Internet Mail Extensions) for file attachments. Obviously transmission times can be longer. An appropriate viewer (hardware or software) is then needed to replay the mail.

An alternative approach has been taken by Olivetti it is an extensions to the PCC video conferencing system. These extensions enable video and voice messages to be left on a PC for viewing by a local or remote user connected over ISDN.

8.2.4. Image Viewing

An image database is considered to be a useful application for attachment to existing databases, or for standalone use. Products have been design for both general use and access via ISDN. Most products are designed not as stand alone image viewers but as part of an information system or kiosk for tourism or specialist applications. An example of the latter is a system from On Demand Information which markets a database for the building industry, accessible over ISDN.

8.2.5. Information Kiosks

Information kiosks are starting to appear with multimedia features. Information kiosks have commercial applications in tourism, government information and education. Extensive multimedia material can be held on local hard discs and updated at regular intervals, manually or via a network connection such as ISDN. The Olivetti PCC video conferencing system allows video conferencing to be incorporated in the design of an information kiosk. All the functions of the system can be programmed from within a high level language such as C++ , Toolbook or Visual Basic.

8.2.6. Distance Learning

Distance learning applications can run on a network file server. They can then be accessed from the local LAN or by ISDN. The high speed of ISDN means that

the remote connection of PC's to LANs is now viable. The response of a distance learning application running over ISDN instead of a local LAN will depend on its design and the amount of visual material. Transfer of audio material will also take time and require a remote workstation set up for audio playback.

A remote access connection over ISDN to a LAN can be set up in three ways:

(a) Setting up a serial synchronous or asynchronous connection using terminal adapters to link the PC into a conventional LAN bridge or router.

(b) Placing an ISDN card in the remote PC and a similar card in a dedicated PC which is also connected to the LAN via an Ethernet card. The latter then acts as a dedicated gateway and is sometimes sold as a dedicated box.

(c) Placing an ISDN card in the PC and a similar card in a server for the LAN

All of these methods require appropriate software to be run on the PC and gateway/server. LAN access is available for the Novell IPX and TCP/IP protocols, for Ethernet and Token Ring LANs.

8.3 Non-Real Time Multimedia Applications

A number of non-real time multimedia applications are available for use over networks which can support audio or video.

8.3.1 File Transfer

For LANs the de facto file transfer standard is the TCP/IP application - FTP. There are various approaches to file transfer on ISDN. Existing file transfer protocols can be used. These may be XModem, and YModem in the case of asynchronous transfers or the Internet FTP as used over LANs and WANS. Either asynchronous communications software such as Crosstalk, Procomm or LAN TCP/IP applications such as FTP are needed or file transfer protocols specifically written for ISDN can be used, using proprietary protocols or the emerging Euro file standard. Data rates of about 1 Mbyte per minute are obtainable on an ISDN2 line.

8.3.2 World Wide Web

Information provision is increasingly moving away from text to graphical interfaces with a multimedia content. The World Wide Web is an example. The World Wide Web has rapidly gained interest through its client implementations such as Mosaic. The delivery of multimedia requires a widespread network capable of delivering at high data rates. Paradoxically, on the Internet, many thousands of users have to compete to use links which are rarely above 2 Mbps in capacity. The interesting feature of ISDN for World Wide Web users is that an ISDN channel offers at least 64 kbps which is dedicated to one user. ISDN in the narrow band form is the most widely available access and delivery medium available. Using remote LAN access products running TCP/IP, it is also relatively easy to implement. ISDN is seen by many in the industry as the ramp through which multimedia networking will gain acceptance. The installed base of ISDN is growing rapidly. ISDN is able to provide connections throughout the world.

There are still major problems of software and hardware compatibility between different ISDN products. The cost of these products is still high but falling. ISDN access to World Wide Web servers on a LAN can be implemented in two ways, routing and bridging.

A Bridge is used to connect two different LAN's that use the same LAN protocol such as Novell IPX or TCP/IP. The bridge acts as an address filter, picking up packets from one LAN and passing on those packets intended for the other LAN. A bridge does not modify the packets or add anything to them. A bridge operates at the data link, Level 2 of the OSI model. A bridge uses the Media Access Control level addresses on LAN adapters to direct packets.

A Router is used to connect two networks that may not use the same LAN protocol. A router uses an inter-networking protocol which is understood by other routers and machines connected to each network. A router operates at the Network, Level 3 of the OSI model. A router uses an addressing scheme such as the Internet address scheme to direct packets. ISDN connected LAN's can use either bridges or routers. Remote access to these LAN's is achieved by enabling the remote workstation to pretend that it is directly connected to the LAN.

Both routers and bridges need to provide basic functions to ensure reliability and security of data. The use of a bridge or router over ISDN also requires some additional mechanisms to reduce the cost of making unnecessary calls on the ISDN. At present this is implemented by 'spoofing' or fooling a LAN that wishes to send these packets to a remote LAN into thinking that these packets have actually been transmitted.

To connect LAN's which are closely coupled and within one organization a bridge has some advantages and may give better performance. To connect LAN's operating between different organizations a router enables more effective management of addressing schemes and security. The most efficient way to interconnect LAN's or remotely attaches a workstation via a dial up connection such as ISDN is by use of the Point to Point Protocol.

8.3.3 Multimedia Mail

Multimedia mail can be delivered via X.400 mail, or Internet mail with MIME (Multipurpose Internet Mail Extensions) extensions for file attachments. Obviously transmission times depend on the network capacity An appropriate viewer (hardware or software) is then needed to view (play) the mail. MIME is a way of transferring multiple objects in a single electronic mail. These objects can be text, images with a JPEG format, 8 bit PCM audio, MPEG video, application specific data, or postscript files.

8.4 Streaming

Streaming is a technology for playing audio and video files (either live or pre-recorded) from a Web page. A user can view the audio or video files directly from the Web server for immediate playback. This avoids time consuming downloads of large files. Streaming may also be referred to as Web casting.

When audio or video is streamed, a small buffer space is created on the user's computer, and data starts downloading into it. As soon as the buffer is full (usually just a matter of seconds), the file starts to play. As the file plays, it uses up information in the buffer, but while it is playing, more data is being

downloaded. As long as the data can be downloaded as fast as it is used up in playback, the file will play smoothly.

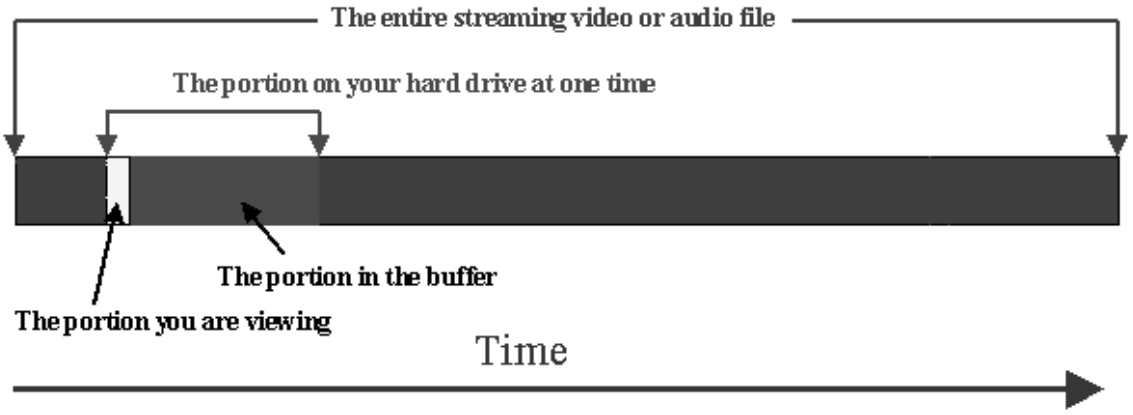


Figure 8.1 Principle of Streaming

Usually there is a delay of only 10-30 seconds before the audio or video starts to play. Streamed files also don't require much bandwidth, so they can be played on computers that use modems to connect to the Internet.

With streaming, users can access lengthy prerecorded audio and video clips to enhance and enrich their study of a topic. Users can also watch or listen to a live event remotely. In the case of distributed learning, streaming audio and video can serve as the primary mode of content delivery.

The people viewing the Web page need to have a player application to view the streamed files--the basic players are free, and are available for Windows, Macintosh, and UNIX computers on the net.

8.4.1 Streaming Audio and Video on the Web

Before the advent of streaming, the delivery of audio and video through the Web was highly impractical. Today, because of streaming, users can access very lengthy pre-recorded audio and video clips to enhance and enrich their study of a topic, or even watch or listen to a live event remotely; in the case of distributed learning, streaming audio and video can even serve as the primary mode of content delivery.

When digital audio and video clips become more than a few seconds or minutes long, the size of the file needed to store the data becomes quite large. If the files are linked or embedded into a Web page, the person browsing your Web site may endure a long delay between clicking on a movie link and seeing that movie play. This is because a large portion of the file--possibly the entire thing--needs to be downloaded before the file can play. Streaming alleviates this wait.

There are two sides to Streaming: what happens on the user's computer, and what happens on the servers.

8.4.2 On User's Computer

When audio or video is streamed, a small buffer space is created on the user's computer, and data starts downloading into it. As soon as the buffer is full (usually just a matter of seconds), the file starts to play. As the file plays, it uses up information in the buffer, but while it is playing, more data is being downloaded. As long as the data can be downloaded as fast as it is used up in playback, the file will play smoothly.

This has some important implications. One is that the media file never exists in its entirety on the user's computer. As data in the buffer is used in playback, it is overwritten by new, incoming data. This means that users never have a complete copy of the file on their computer, so every time they want to view the file, they have to come back to your Web site. Also, jumping forward or backward into the file requires the buffer be reloaded with pertinent data, with the subsequent delay of several seconds. This may hurt interactivity.

However streaming does make it possible to watch or hear large media files on computers that don't have much free hard-disk space (since only a portion of the file is on the computer at any one time).

8.4.3 On Servers

Two servers are required to deliver streaming media. First, the user requests a normal Web page (.html) that contains links to streaming media. This Web page can reside on any Web server. The link the user clicks on actually links to an

intermediate file (.ram in the case of Real media). The .ram file is just a plain text file containing nothing more than the address of the streaming media file. The job of the .ram file is to direct the request to the streaming server.

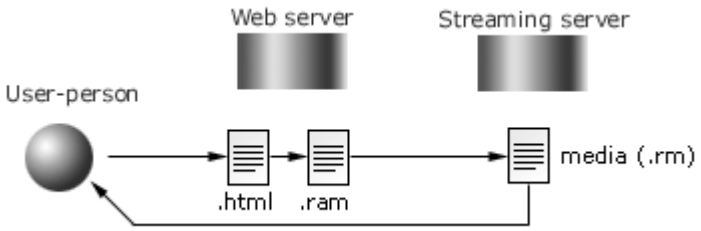


Figure 8.2 Roles of Servers in Streaming

The streaming server houses the media file itself (such as a .rm file for Real media). The streaming server's job is to balance the load of delivering streaming files through the network to many users.

8.4.4 Technologies available for streaming

Three are many ways to deliver streaming audio and video, and no clear standard has yet emerged. Because of this, you should be aware of the fact that the landscape is changing very rapidly right now. To make it even more confusing, different streaming media viewers are starting to handle many different file formats.

The following are some of the several streaming media formats .

- Real Networks' Real Media format (.rm file extension)
- Microsoft's Windows Media format (.wmv file extension)
- Apple Computer's QuickTime format (.mov file extension)
- Coming: MPEG format (.mpg file extension)

So how do you choose which to use? Your decision will probably be based on which type of server you have available to you, what your users may already have installed on their computers, and what software and hardware you have available to develop the media. (For example, encoding Real streaming media

seems to work much better with a PC than with a Mac, but QuickTime streaming is easier to build right now using a Mac.)

However, the format in which you prepare the file no longer dictates what software you might use to view it. As mentioned before, the landscape is changing rapidly.

8.4.5 Streaming Media Viewers

The three major players at the moment appear to be

- Real Networks' Real Player,
- Microsoft's Windows Media Player, and
- Apple Computer's QuickTime Player.

Each one is a freely-downloadable player. All are available for both Macintosh and Windows, and Real is also available for UNIX. In addition, each requires a server to reap the full benefits of streaming. Each can be presented on a Web page either through embedding or linking

8.4.6 SMIL (Synchronized Multimedia Integration Language)

SMIL is a set of tags that are very similar to the tags used in HTML. A SMIL file can be created with a text editor (by typing tags) or with a utility program, just like web pages can be created by typing tags or with a web page editor. SMIL is at an earlier stage of development that HTML, so the number of utility programs is small, and for most applications, typing tags is the only way to proceed.

8.4.7 Combining streams with SMIL

Streaming can be used for more than just delivering a single channel of audio or video. Using SMIL (synchronized multimedia integration language), multiple streams can be combined into one presentation.

Suppose you want to display text at a specific time during an music clip in order to make a point about what's happening in the music. Or, suppose you want to display a video of a speaker on one side of the screen, and coordinate his Power

point slides with his talk on the other side of the screen. SMIL is used to combine these streams, as well as build the graphic interface to present them.

Let's take a look at some examples of how SMIL can be used to combine multiple streams of media:

- audio + text
- audio + graphics
- video + graphics
- video + text
- graphics + text + audio

8.4.8 Role of Servers in SMIL

The user requests the web page from a normal web server as before. The web page contains a link to a .ram file, which also lives on the web server. The Ram file points to a SMIL file that lives on the streaming server. The job of the SMIL file is to divide the streaming window into areas, then coordinate different streams of content in each area together over time. So, in the case of that speaker and his Power point slides, a SMIL file might divide the streaming window into three areas: a banner graphic at the top, a space for the video of the speaker on the left, and a space for the slides on the right. The SMIL file also contains time codes to describe exactly when each piece of media appears.

The media files called by the SMIL file are often located on the streaming server, but they don't have to be. It is possible to refer to files located on other servers. The risk of doing this is that slow network response might throw the presentation's timing off if the media is not delivered in a timely manner.

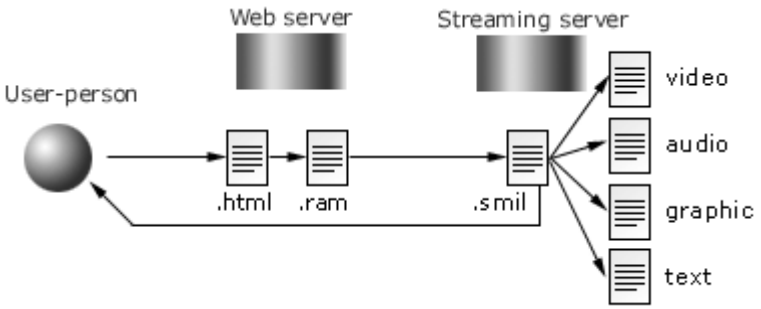


Figure 8.3 Roles of Servers in SMIL

8.4.9 Live streaming

Remember that in streaming, the entire file never exists on the user's computer at any one time. Rather, the user receives it in a steady stream and views it as it comes in. So, what if you could build that stream very quickly, as fast as your video camera or microphone could record the event, and send that stream out to the users as soon as it was ready, rather than saving it to a file? In live streaming, that's exactly what happens. The result is that users can view an event almost at it occurs; in most cases, there is only a 5-10 second delay from the time the moment occurs to the time that moment is delivered to the user. This is used in radio broadcasting and video news on net.

8.4.10 Roles of Servers in Live Streaming

First, a video camera or microphone capturing the live event is connected to a separate encoding computer. The encoding computer very quickly processes the video or audio signal and sends it in a continuous flow to the streaming server. The streaming server simply broadcasts that incoming signal to anyone who's connected.

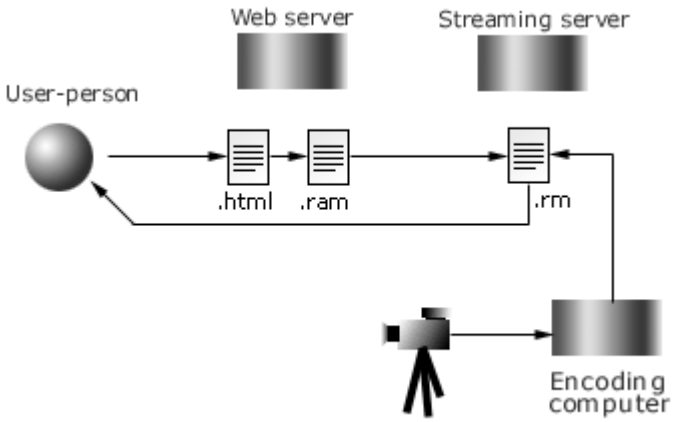


Figure 8.4 Roles of Servers in Live Streaming

In the diagram above, it's important to note that the .rm file is not a complete file containing the entire event; instead, it contains whatever chunk of data is currently being streamed to the users at that point in time.

So this time, the server never holds a file containing the entire event. As an option, you could choose to archive the event on the encoding machine while it's being sent to the server; if you did, you would have a complete .rm file on the encoding machine once the event was over. You would then move that archived file to the streaming server for viewers to watch later.

8.5 RTP (Real-time Transport Protocol)

RTP (Real-time Transport Protocol) is the Internet-standard protocol for the transport of real-time data, including audio and video. It can be used for media-on-demand as well as interactive services such as Internet telephony. RTP provides end-to-end network transport functions but does not address resource reservation and does not guarantee quality-of- service for real-time services. The data transport is augmented by a control protocol (RTCP) to allow monitoring of the data delivery in a manner scalable to large multicast networks, and to provide minimal control and identification functionality. RTP and RTCP are designed to be independent of the underlying transport and network layers.

It was originally designed as a multicast protocol, but has since been applied in many unicast applications. It is frequently used in streaming media systems as well as videoconferencing and push to talk systems, making it the technical foundation of the Voice over IP industry. It goes along with the RTCP and it's built on top of the User Datagram Protocol (UDP). Applications using RTP are less sensitive to packet loss, but typically very sensitive to delays, so UDP is a better choice than TCP for such applications.

The services provided by RTP include:

- Payload-type identification - Indication of what kind of content is being carried
- Sequence numbering - PDU sequence number
- Time stamping - presentation time of the content being carried in the PDU
- Delivery monitoring

The protocols themselves do not provide mechanisms to ensure timely delivery. They also do not give any Quality of Service (QoS) guarantees. These things have to be provided by some other mechanism.

Also, out of order delivery is still possible, and flow and congestion control are not supported directly. However, the protocols do deliver the necessary data to the application to make sure it can put the received packets in the correct order. Also, RTCP provides information about reception quality which the application can use to make local adjustments. For example if congestion is forming, the application could decide to lower the data rate.

8.5.1 Packet Structure of RTP

The Packet structure of RTP is shown below-

+ Bits	**0-1**	**2**	**3**	**4-7**	**8**	**9-15**	**16-31**
0	Ver	P	X	CC	M	PT	Sequence Number
32	Timestamp						
64	SSRC identifier						
96	CSRC identifiers						
96+(CC×32)	Additional header (optional), indicates length “AHL”						

96+(CC×32) + (X×(AHL+16))	Data

Ver (2 bits) indicates the version of the protocol. Current version is 2. P (one bit) is used to indicate if there are extra padding bytes at the end of the RTP packet. X (one bit) indicates if the extensions to the protocol are being used in the packet. CC (four bits) contains the number of CSRC (contributing source) identifiers that follow the fixed header. M (one bit) is used at the application level and is defined by a profile. If it's set, it means that the current data has some special relevance for the application. PT (7 bits) indicates the format of the payload and determines its interpretation by the application. SSRC indicates the synchronization source. The sequence number (16 bits) increments by one for each RTP data packet sent, and may be used by the receiver to detect packet loss and to restore packet sequence. The timestamp (32 bits) reflects the sampling instant of the first octet in the RTP data packet. The SSRC field (32 bits) identifies the synchronization source. This identifier is chosen randomly, with the intent that no two synchronization sources within the same RTP session will have the same SSRC identifier. The CSRC list (0 to 15 items, 32 bits each) identifies the contributing sources for the payload contained in this packet. The number of identifiers is given by the CC field. If there are more than 15 contributing sources, only 15 may be identified.

8.6 Summary

Multimedia networking applications are playing amazing roles in real life. One will experience its use very shortly, if he has not experienced it. Streaming has made possible transfer of audio and video information on web, which was highly impracticable before. Now students can get stored or live lectures on various topics by sitting in front of computer through internet. RTP has become standard of real time transfer of multimedia transfer.

8.7 Keywords

SMDS- SMDS, which stands for Switched Multi-megabit Data Services, was a connectionless service used to connect LANs, MANs and WANs to exchange data. SMDS was based on the IEEE 802.6 DQDB standard. SMDS fragmented its datagrams into smaller "cells" for transport, and can be viewed as a technological precursor of ATM.

H.320- H.320 is an umbrella recommendation by the ITU-T for running Multimedia (Audio/Video/Data) over ISDN based networks. The main protocols in this suite are H.221, H.230, H.242, audio codecs such as G.711 and G.723, and video codecs such as H.261 and H.263.

X.400- X.400 is a suite of ITU-T Recommendations that define standards for Data Communication Networks for Message Handling Systems (MHS).

MIME- Multipurpose Internet Mail Extensions (MIME) is an Internet Standard that extends the format of e-mail to support text in character sets other than US-ASCII, non-text attachments, multi-part message bodies, and header information in non-ASCII character sets. Virtually all human-written Internet e-mail and a fairly large proportion of automated e-mail is transmitted via SMTP in MIME format.

T.120- T.120 is an ITU-T recommendation that describes a series of communication and application protocols and services that provide support for real-time, multipoint data communications. It is used by products such as Microsoft NetMeeting and Lotus Sometime to support application sharing, real-time text conferencing and other functions.

8.8 Self Assessment Questions

Q1. Explain the different real time applications of multimedia networking.

Q2. What are the non real time applications available on the network, which support audio and video?

Q3. What is SMIL? Explain the role of server in SMIL.

Q4. What is RTP? Explain the packet structure of RTP.

Q5. What is streaming? How streaming is useful in multimedia networking?

Q6. Explain Live Streaming in detail.

Q7. Explain the different streaming technologies and media viewers?

8.9 References

1. Internet Book, Third Edition, Dougles E. Commer, Pearson Education.

2. Multimedia Communications: Applications, Networks, Protocols and Standards, Fred Halsall, Addison-Wesley .

3. Multimedia Communication Systems: Techniques, Standards, and Networks, K. R. Rao, Zoran S. Bojkovic, Dragorad A. Milovanovic, Prentice Hall.