1.1 Computer Network

Network is the collection of computer, software and hardware that are all connected to each other to help their work together. A network connects computers by means of cabling system (or wireless media), specialized software and devices that manage data traffic. A network enables users to share files and resources such as printer as well as send message electrically to each other.

Computer network falls under two types

1) Client server network
2) P2P (peer-to-peer) network

- **Client Server Network**

Each client is assigned as account name and password that is verified by an authentication service. The authentication service guards access to the network. With the centralization of user accounts, security and access control, server based networks simplify the administration of large network.

The concentration of network resources such as files, printers and applications on servers also makes it easier to backup and maintain the data. Resource can be located on specialized dedicated servers for easier access.

Advantages

- Easier to administer when the network is large.
- All data can be backed up on one central location.

Disadvantages

- Requires expensive, more powerful hardware for the server machines.
- Has a single point of failure user data is unavailable when the server is down.
- Requires expensive specialized network administrative and operational software.
- Requires a professional administrator.

- **Peer-to-Peer Network**

Network computers act as equal partners, or peers. Each computer can take on the client function or the server function.

Suppose computer A may request for a file from computer B, which then sends file to computer A. In this case, computer A acts like the client and computer B as server.

At a later time, their role may be reserved; individual users control their own resources.

The users may decide to share certain files with other users. The users may also require passwords before they allow others to access their resources. Since individual users make these decisions, there is no central point of control or administration in the network.

When a computer acts as a server, the user of that machine may experience reduced performance as the machine server the requests made by other system.

Advantages

- Less expensive to implement.

Downloaded from CSIT Tutor

- Doesn't require additional specialized network administration software.
- Doesn't require a dedicated network administrator.

Disadvantages

- Less secure.
- Doesn't scale well to large networks, and administration becomes unmanageable.
- Each must be trained to perform administrative tasks.
- All machines sharing resources negatively impact the performance.

Advantages and Disadvantages of Computer Network

- **Advantages**

1. File Sharing

The major advantage of a computer network is that it allows file sharing and remote file access. A person sitting at one work station that is connected to a network can easily see files present on another workstation provided he/she is authorized to do so.

If the files are stored on server and all of its clients share that storage capacity, then it becomes easier to make a file available to multiple users.

2. Resource Sharing

For example, if there are twelve examples in an organization, each having their own computers, they will require twelve modems and twelve printers if they go to use resources at the same time. A computer network on the other hand provides cheaper alternative by the provision of resources sharing. All computers can be interconnected using a network and just one modem and printer can efficiently provides the services to all twelve users.

3. Inexpensive Set-Up

Shared resources mean reduction in hardware costs. Shared files means reduction in memory requirements, which indirectly means reduction in file storage expenses.

4. Flexible Handling

A user can log on to the computer anywhere on the network and access his/her files. This offers flexibility to the user as to where he/she should be during the course of his/her routine.

5. Increased Storage Capacity

A standalone computer might fall short of storage memory, but when many computers are on a network the memory of different computers can be used in such a case.

- **Disadvantages**

1. Security

If a computer is on a network, a hacker can get unauthorized access by using different tools. In case of big organizations, various security software need to be used to prevent theft of any confidential and classified data.

2. Virus Attack

If even one computer on a network gets affected by a virus, there is a possible threat for the other systems getting affected too. Viruses can spread on a network easily, because of inter-connectivity of workstations.

Downloaded from CSIT Tutor

3. **Lack Of Robustness**

If the main file server of computer network breaks down, the entire system becomes useless.

4. **Need Of Expert Handler**

The technical skills and know-how required to operate and administer a computer network is considerably high.

5. **Lack Of Independence**

Since most computers have a centralized server and dependent clients, the clients/users lack any freedom whatsoever.

Network Edge

- **End System**

→ End system are also referred to as host because they host (ie, run) application program such as a web browser program, a web server program, an email reader program.

→ Host are further divided into two categories:

- Clients
- Servers

→ Informally, clients tend to be desktop and mobile pc's and so on, whereas servers tend to be more powerful machine that stores and distribute web pages, stream video so on.

- **Clients and Servers**

→ A client is a program running on one end system that requests and receive a service from a server running on other end system.

→ Not all internet applications are client-server model; they are also peer-to-peer model like Bit Torrent and eMute.

Network Core

→ It means the approach to moving data through a network of links and switches.

→ There are two types:

- Circuit Switching
- Packet Switching

→ In **circuit switched network**, the resources needed along a path (buffer, link transmission rate) to provide for communication between the end systems are reserved for the duration of the communication session between the end systems.

→ It reserves a constant transmission rate in the network's links for the duration of the communication. Since bandwidth has been reserved for this sender-to-receiver. Connection, the sender can transfer the data to the receiver at the guaranteed constant rate.

→ In **packed switched network**, the resources are not reserved for a sessions message use the resources on demand, and as a consequence may have to wait (i.e. queue) for access to a communication link.

Downloaded from CSIT Tutor

→ The packet is sent into the network without reserving and bandwidth if one of the links is busy because other packets need to be transmitted over the link at the same time, our packet will have to wait in a buffer at the sending side of the transmission link, and suffer a delay.

There are two approaches in packet switched network

1. **Datagram Network**
   - Any network that forwards the packets according to the destination address is called a datagram network.
   - The routers in the internet forwards packets according to the destination address. Hence, internet is datagram network.
2. **Virtual Circuit Network**
   - Any network that forwards the packets according to virtual circuit identifier (fixed route) is called a virtual circuit network.
   - Preplanned route established before packets sent.
   - Examples are X25, Frame relay, ATM technologies.

Network Access

1. Dial-Up

→ Accessing the internet over ordinary analog telephone lines using a dial-up modem.

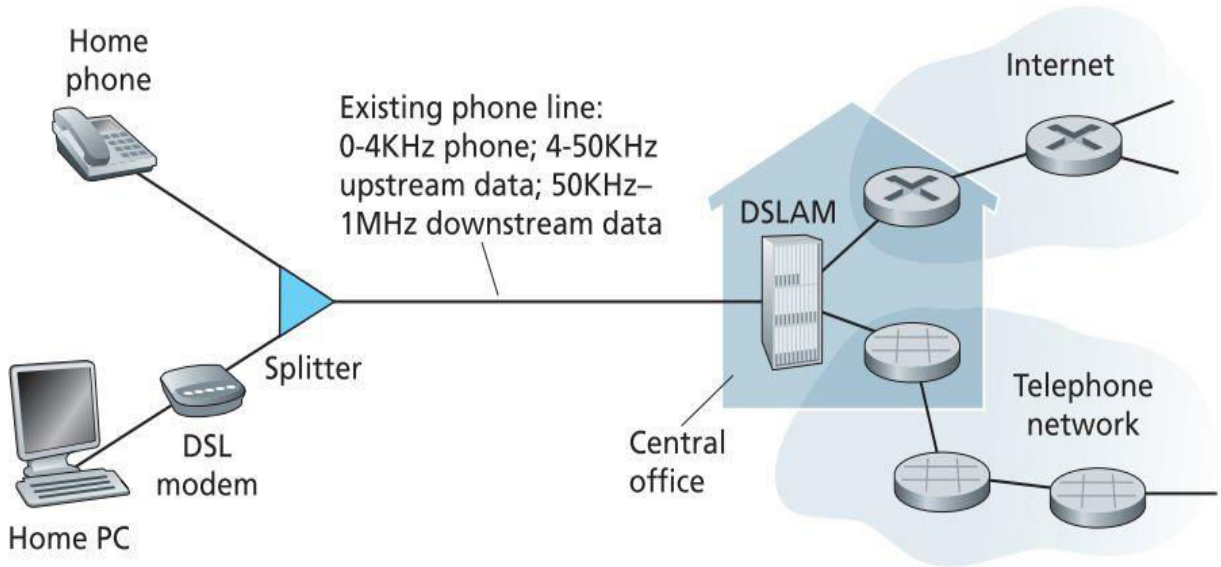
→ The term "dial-up" is used because the user software actually dials an ISP's phone number and makes a traditional phone connection with the ISP.

→ Two major drawbacks:

- Extremely slow, maximum 50 kbps.
- Cannot use phone line while accessing internet.

2. DSL (Digital Subscriber Line)

→ DSL internet is accessed from the same company that provides it wired local phone access.



→ Advantages:

Downloaded from CSIT Tutor

a) High data rate (download 1-2 mbps, upload 128 kbps-1mbps)
b) Users can simultaneously talk on the phone and access internet.

3) Cable

→ Cable internet access make the use of cable TV company's existing infrastructure

→ Both the fiber and co-axial cable are employed, it is also called hybrid fiber coax (HFC)

→ Requires cable modem, which connects to the home PC through Ethernet port.

4) Wireless Network

→ Wireless LAN (Wi-Fi)

IEEE 802.11

→ Wide Area Wireless Access e.g.:- 3G, GPRS.

→ WiMAX

→ Intel WiMAX 2009

→ k/a IEEE 802.16 is a long distance derivation of the 802.11 WiFi protocol speed 5-10 Mbps.

Physical Media

1. Guided Media

The waves are guided along a solid medium.

a) Twisted Pair Cable
b) Coaxial Cable
c) Optical Fiber

a) Twisted Pair Cable

It consists of two insulated strands (each about 1mm thick) of copper wire twisted around each other to form a pair. One or more twisted pairs are used in it. The purpose of twisting is to eliminate electrical interference from the wires and cancels any noises from the adjacent pair. The more twist per linear foot, the greater the effect.

There are two types:

i. SIP (Shielded Twisted Pair)

- Has a foil or a wire braid wrapped around the individual wires of the pair.
- Minimizes EMI radiation.

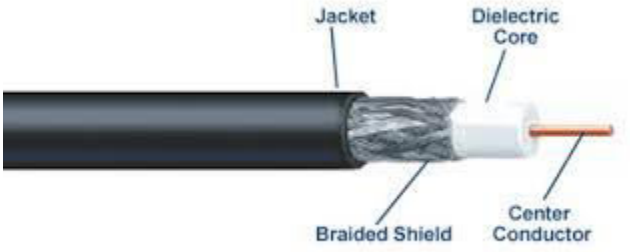
ii. UTP (Unshielded Twisted Pair)

No shielding and is more victim of EMI but is the most frequently used because it is inexpensive and easier to install.

b) Coaxial Cable

It is made of two conductors that share the same axis, the centre is a copper wire that is insulated by a plastic coating and then wrapped with an outer conductor (usually a wire braid).

Downloaded from CSIT Tutor



There are two types:

i. **Thick Net**

→ 0.38 inch in diameter

→ Also called 10b5 cable, which means speed 10 and signal can be strong up to 500m.

ii. **Thin Net**

→ 0.25 inch in diameter.

→ Similar to the material commonly used in cable TV.

→ Also called 10b2 cable, means 10 mbps and can carry up to 200 m before being weak.

c) Optical Fiber

- A thin, flexible medium that conducts pulses of light, with each pulse representing a bit.
- Can support high bit rates, up to Gbps, immune to EMI, has very low signal attenuation up to 100 km
- Standard is optical carrier (OC), ranges from 51.8 Mbps to39.8 Gbps.
- OC-1, OC-3, OC-12, OC-24, OC-48, OC-96, OC-192, OC-768 are available.
- OC-n, where the link speed equals n*51.8 Mbps.

There are two types:

i. **Single Mode (Mono Mode)**

For longer distance and LASER is used as light source.

ii. **Multi Mode**

For shorter distance and LED is used as light source.

2. Unguided Media

The waves propagate in the atmosphere and in outer space, such as in a wireless LAN or a digital satellite channel.

a) Terrestrial Radio Channel

b) Satellite Radio Channel

a) Terrestrial Radio Channel

→ Carry signals in the electromagnetic spectrum.

→ No need of physical wire to be installed, can penetrate walls, provide connectivity to mobile user, and potentially carry a signal for long distances.

Downloaded from CSIT Tutor

→ May face
- Path loss
- Shadow fading(which decrease the signal strength as signal travels over a distance and around/through obstructing objects)
- Multiple fading (due to signal reflection off of interfering objects)

→ E.g. wireless LAN, cellular access technology.

b) Satellite Radio Channels

→ Links two or more Earth-based microwave transmitter/receiver known as ground stations.

→ Satellite receives transmissions on one frequency band, regenerates the signal using a repeater, and transmits the signal on another frequency.

→ Two types:

i. Geostationary Satellite
- Permanently remain above the same spot on Earth at 36,000 km above surface.
- Propagate delay of 280 ms.

ii. Low Earth Orbiting (LEO) satellite
- Placed much closer to Earth and do not remain permanently above one spot on Earth.
- Rotate around Earth and may communicate with each other as well as with ground stations.
- For continuous coverage, many satellites need to be placed.

Downloaded from CSIT Tutor

1.2 Protocol Layers

Protocol

In computer science and telecommunications, a communication protocol is a system of digital rules for data exchange within or between computers. When data is exchanged through computer network, the rule system is called a network protocol.

Protocol Layer

In software and hardware environment of two or more communication devices or computers in which a particular network protocol operates. A network protocol may be thought of as a set of more or less independent protocols, each in a different layer or level. The lowest layer governs direct host-to-host communication between the hardware at different hosts; the highest consists of user application programs. Each layer uses the layer beneath it and provides a service for the layer above. Each networking components hardware or software on one host uses protocols appropriate to its layer to communicate with the corresponding component on another host.

Layered Architecture

A reference model (OSI and TCP/IP) is a conceptual rule of how communications should take place. It addresses all the processes required for effective communication and divides these processes into logical groupings called layers. When a communication system is designed in this manner, it is known as layered architecture.

Need of Layered Architecture

→ It divides the network communication process into smaller and simpler components, thus aiding component development, design and troubleshooting.

→ It allows multi-vendor development through standardization of network components.

→ It allows different various types of network software and hardware to communicate.

OSI Model (Open System Inter Connection)

Layer 7	Application	• Provides a user interface
Layer 6	Presentation	• Presents data • Handles processing such as encryption/decryption

Downloaded from CSIT Tutor

Layer 5	Session	• Keeps different applications data separate
Layer 4	Transport	• Provides reliable or unreliable delivery • Performs error correction before retransmit
Layer 3	Network	• Provides logical addressing, which routes use for path determination
Layer 2	Data Link	• Combines packets into bytes provides access to media using MAC address performs error detection, not correction.
Layer 1	Physical	• Specifies voltage wire speed and pin-out of cables.

TCP/IP (Transmission Control Protocol/IP) Model

Application	Process (FTP, SMTP
Transport	Transport
Internet	Internet
Network Access	Network Access

OSI Layers

Layer 7	Application
Layer 6	Presentation
Layer 5	Session
Layer 4	Transport
Layer 3	Network
Layer 2	Data Link
Layer 1	Physical

Layer 7: Application

This is the layer that actually interacts with the operating system or application whenever the uses to transfer files, read messages or perform other network-related activities.

Layer 6: Presentation

Layer 6 takes the data provided by the Application layer and converts it into standard format that other layers can understand.

Layer 5: Session

Layer 5 establishes, maintains and ends communication with the receiving device.

Downloaded from CSIT Tutor

Layer 4: Transport

This layer maintains flow control of data and provides for error-checking and recovery of data between the devices. Flow control means that the Transport layer looks to see if data is coming from more than one application and integrates each applications data into a single stream for the physical network.

Layer 3: network

The way that the data will be sent to the recipient device is determined in this layer. Logical protocols, routing and addressing are handled here.

Layer 2: Data Link

In this layer, the appropriate physical protocol is assigned to data. Also, the type of network and the packet sequencing is defined.

Layer 1: Physical

This is the level of the actual hardware. It defines the physical characteristics of the network such as connections, voltage levels and timing.

TCP/IP Layer

Layer 4: Application Layer

This layer defines TCP/IP application protocols and how host programs interface with Transport layer services use the network.

Layer 3: Transport Layer

The purpose of this device is to permit devices on the source and destination hosts to carry on a conversation. It defines the level of service and status of the connection used when transporting data.

Layer 2: Internet Layer

This layer packs data into data packets known as IP datagram, which contain source and destination address (IP address) information that is used to forward the datagram between hosts and across networks.

Layer 1: Network Access Layer

This layer defines details of how data is physically sent or optically signaled by hardware devices that interface directly with a network medium, such as co-axial cable, optical fiber or twisted pair, copper wire.

Data Encapsulation and De-encapsulation

- When a host transmits data across a network to another device, the data goes through encapsulation.
- At the receiving device, the data goes through de-encapsulation.

Layer Addressing

Downloaded from CSIT Tutor

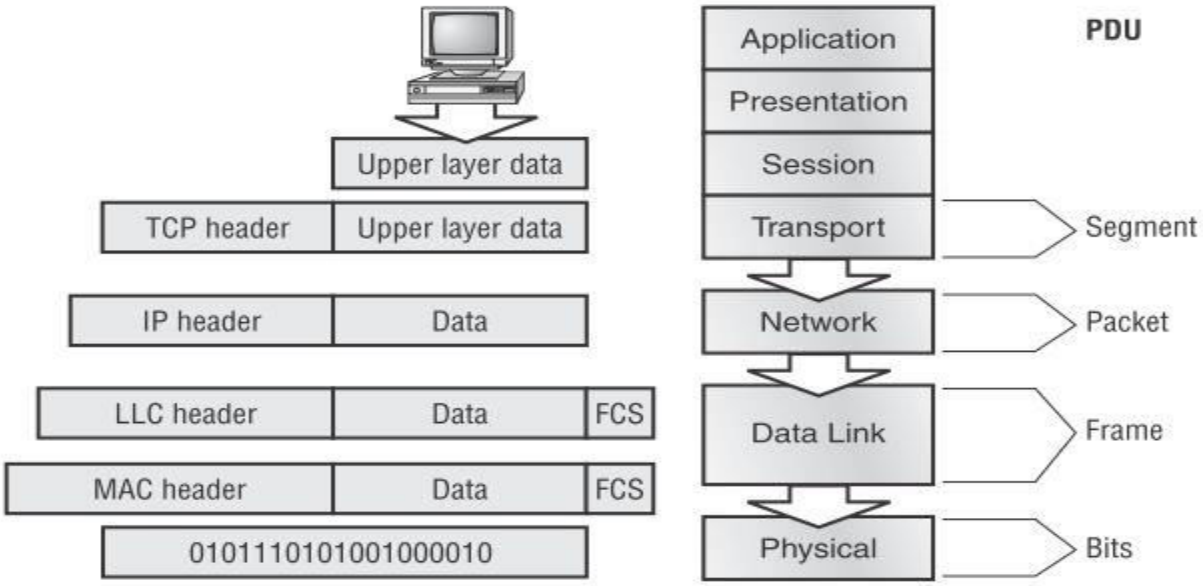


Fig: Data Encapsulation

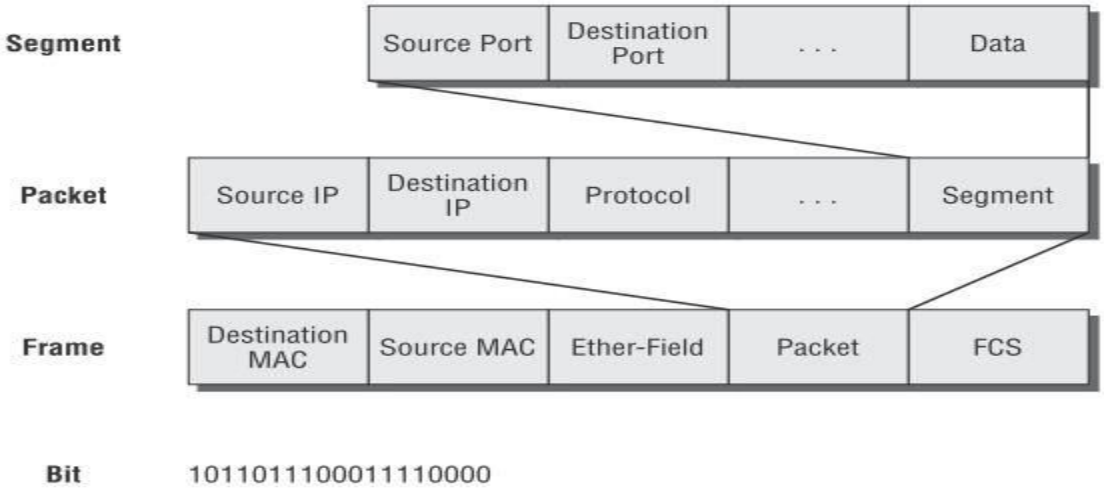


Fig: PDU and Layered Addressing

At a transmitting side, the data encapsulation method works like this:

1. User information is converted to data for transmission on the network.
2. Data is converted to segments and a reliable connection is set up between the transmitting and receiving hosts.
3. Segments are converted to packets or data grams, and a logical address is placed in the header so each packet can be routed through the internetwork.
4. Packets or datagram are converted to frames for transmission on the local network. Hardware (Ethernet) addresses are used to uniquely identify hosts on a local network segment.
5. Frames are converted to bits, and a digital encoding and clocking scheme is used.

Downloaded from CSIT Tutor

Network Entities

Hub

→ It is a device for connecting multiple devices together and making them act as a single network segment.

→ It has multiple I/O ports, in which a signal introduced at the input of any port appears at the output of every port except the original incoming.

→ Doesn't examine or manage any or the traffic that comes through it, any packet entering any port is rebroadcast on all other ports.

Switch

→ Device used to connect devices together on a computer network.

→ A switch is considered more advanced than hub because a switch will only send a message to the device that need or requests it rather than broadcasting the same message out of each of its ports.

Router

→ A device that forwards data packets between computer network creating an overlay internetwork.

→ It is connected to two or more data links from different network.

→ When a data packet comes in one of the lines, the router reads the address information in the packet to determine its ultimate destination.

Bridge

→ It is a network device which connects two or more LANs.

Repeater

→ It is an electronic device that receives a signal and retransmits it at a higher level or higher power, and onto the other side of an abstraction, so that the signal can cover long distances.

→ In telecommunication, the term repeater has the following standardized meanings.

1. An analog device that amplifies as input signal regardless of its nature.
2. A digital device that amplifies, reshapes, retimes or performs a combination of any of these functions on a digital input signal for retransmission.

Downloaded from CSIT Tutor

1.3 Application Layer

Web: HTTP	File Transfer: FTP	E-mail: SMPT, POP3, IMAP	Remote login: TELNET	Network Management: SNMP, NFS, TFTP	Name Management: DNS

The Web and HTTP

Hypertext transfer protocol (HTTP) works with the world wide web (WWW) which is the fastest growing and most used part of the internet. It is popular because of the ease with which it allows access to information. A web browser is a client-server application, which means that it requires both a client and a server component in order to function. A web browser presents data in multimedia formats on the web pages that use text, graphics, sound and video. The web pages are created with a format language called hypertext Markup language (HTML). HTML directs a web browser on a particular web page to produce the appearance of the page in a specific manner. In addition HTML specifies locations for the placement of text, files and objects that are to be transferred from the web server to the web browser.

Hyper links make the World Wide Web easy to navigate. A hyper link is a object, world phrase or picture on a webpage. When that hyperlink is clicked it directs the browser to a new webpage. The webpage contains an address location known as a uniform resource locator (URL).

In the URL, http://www.ekantipur.com/np/pictures, here the http:// tells the browser which protocol we use. The second part "www" is the host name or a name of a specific machine with a specific IP address. The last part "/pictures/" identifies the specific folder, location on the server that contains the default web page.

A web browser usually opens to a starting or home page. The URL of the homepage, has already been stored in a configuration are of the web browser and can be changed at any time. From the starting page, click on one of the webpage hyperlinks or type a URL in the address bar of the browser. The web browser examines the protocol to determine, if it needs to open the other program and then determines the IP address of the web browser using DNS. Then transport layer, network layer, data link layer and physical layer work together to initiate a session with the server contains the folder name of the webpage location. The data can also contain a specific file name for HTML page. If no name is given then the default name as specified in the configuration on the server is used.

The server response to the request by sending to the web client all of the text, audio, video and graphic files specified in the HTML instructions. The client browser reassembles all the files to

Downloaded from CSIT Tutor

create a view of the webpage and then terminates the session. If another page that is located on the same or different server is clicked the whole process begins again.

HTTP Message format

i) HTTP request message format

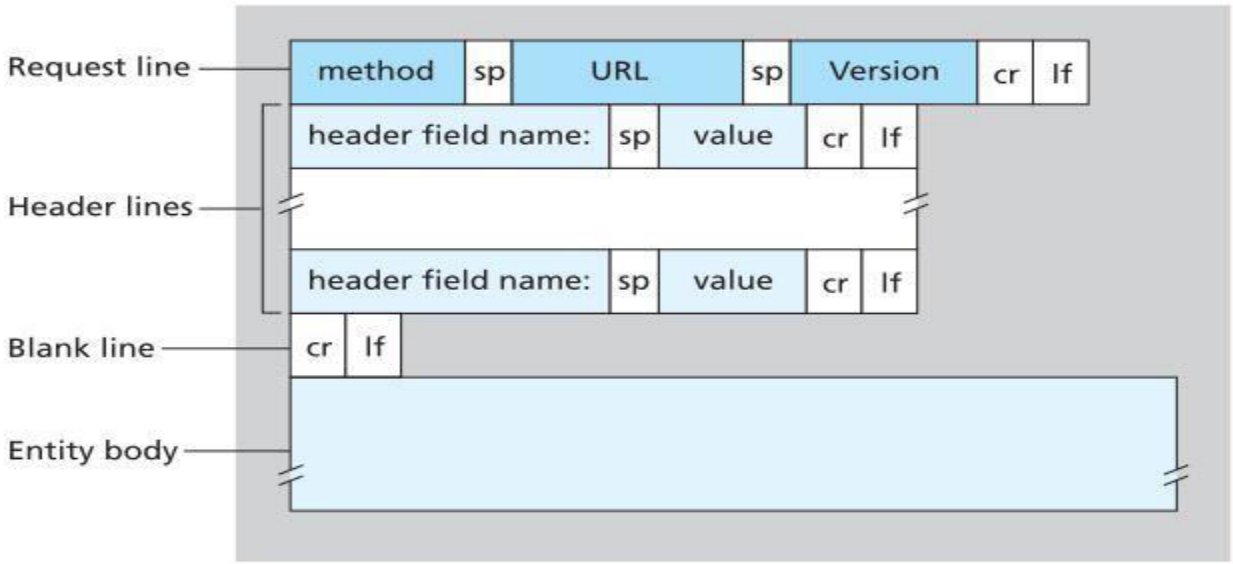


Fig: General format of an HTTP request message

Below we provide a typical HTTP request message:

GET: /somedir/page.html HTTP/1.1
HOST: www.espnsoccernet.com
Connection: close
User-agent: Mozilla/4.0
Accept-language: fr

→ Message consists of five lines (may be more), each followed by a carriage return (cr) and line feed (lf)

→ First line is called request line; the subsequent lines are called the header lines.

→ The request line has 3 fields.

i) The method field

- GET –to browse a particular website
- POST –to search with keywords (entity body is not empty)
- HEAD –requests a HTTP message but leaves out the requested object. Application developers
- PUT –used with web publishing tools to upload objects.
- DELETE –to delete an object on a web server.

ii) URL field

- -/somedir/page.html.

Downloaded from CSIT Tutor

iii) the HTTP version field

- -HTTP/1.1 is the version 1.1 of HTTP.

→ Let's look at the header lines

- Host: www.espnsoccernet.com specifies the host on which the object resides.
- Connection: close is telling the server to close the connection after sending the requesting object.
- User-agent: specifies the browser type.
- Accept-language: fr indicates that the user prefers to receive. French version of the object, if exists on the server, otherwise the server should sent its default version.

ii) HTTP Response Message Format

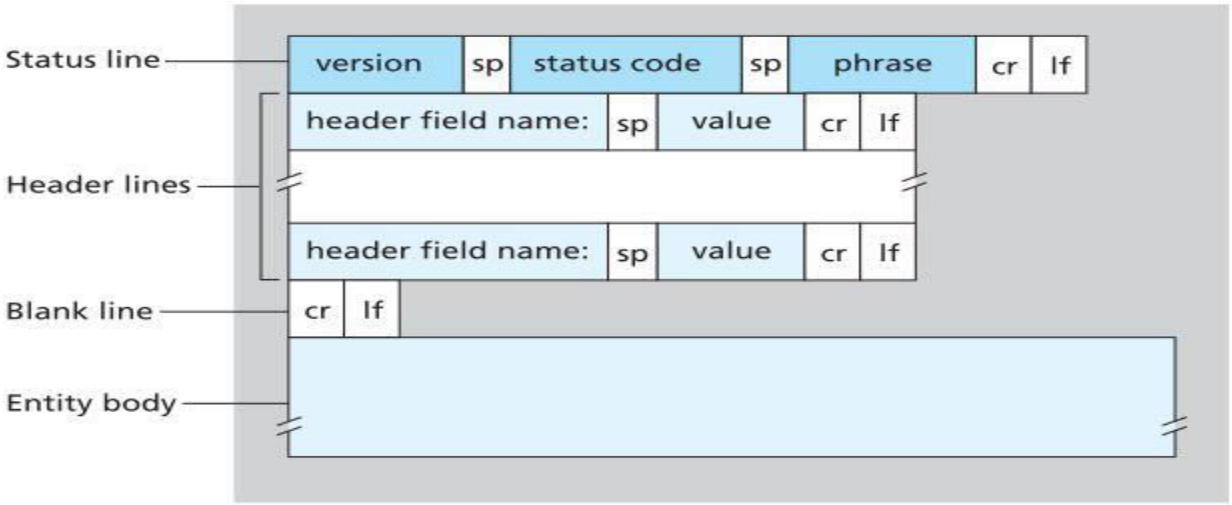


Fig: General format of HTTP response message.

Consists of two parts:

i) Status line
ii) Header Lines

The status line has 3 sections

- The protocol version field.
- The status code.
- Corresponding status message.

A few details about status code and their phrases

200 OK – request succeeded and the information is returned in response.

301 Moved permanently – requested object has been permanently moved, the new URL specified in location: header of the response message.

400 Bad Request – This is a generic code indicating that the request could not be understood by the server.

404 Not Found – The requested document doesn't exist on the server.

Downloaded from CSIT Tutor

505 HTTP Version Not Supported – The requested HTTP protocol version is not supported by the server.

Example:
HTTP/1.1 200 OK
Connection: close
Date: Sun, 09 Mar 2014 09:45 GMT
Server: Apache/1.3.0 (UNIX)
Last–modified: Fri, 6 Jan 2014 08:00:15 GMT
Content-length: 6978
Content-Type: text/html

Connection: Close
— To tell the client that it is going to close the TCP connection after sending the message.
Date:
— Indicates the time and date when the HTTP response was created and sent by the server.
Server:
— Indicates that the message was generated by an Apache web server.
Last-modified:
— Indicates the time and date when the object was created or last modified.
Content-Length:
— Indicates the number of bytes in the object being sent.
Content-Type:
— Indicates that the object in the entity body is HTML text.

Cookies

→ As HTTP server is stateless, web servers can handle thousands of simultaneous TCP connections.
→ It is often desirable for a website to identify users, either because the server wishes to restrict user access or because it wants to serve content as a function of the user identify.
→ For the purpose, HTTP user cookies, which allow sites to keep track of users.
→ Cookie technology has four components.
- A cookie header line in the HTTP response message.
- A cookie header line in the HTTP request message.
- A cookie file kept on the user's end system and managed by the user's browser.
- A back-end database at the website.

→ In the HTTP response message,
→ Set cookie : 1783
→ In the HTTP request message.
Cookie: 1783

Downloaded from CSIT Tutor

FTP (File Transfer Protocol)

FTP is reliable, connection-oriented service that uses TCP to transfer files between systems that support FTP. The main purpose of FTP is to transfer files form one computer to another by copying and moving files from servers to clients, and from clients to servers. When files are copied from a server, FTP first establishes a Control Connection between the client and the server. Then, a second connection is established, which is a link between the computers through which data is transferred. Data transfer can occur in ASCII mode or in binary mode. These modes determine the encoding used for data file, which in the OSI model is a presentation layer task. After the file transfer has ended, the data connection terminates automatically when the entire session of copying and moving files is complete, the command link is closed when the user logs off and ends the session.

Fig: Control and data connections

Client

- Initiates a control TCP connection.
- Sends the user identification and password over the control connection.
- Also sends over the control connection, commands to change the remote directory.

Server

- When a request receives, FTP server starts TCP data connection.
- Sends exactly one file over the data connection and then closes the data connection.
- FTP data connection opens again for another data transfer.
- Non-persistent connection.

FTP Commands and Replies

→ Each successive command follows CR and LF.

→ Each command consists of fur uppercase ASCII characters, some with optional arguments.

→ Some of the more common commands are given below :

- USER username: used to send the user identification to the server.
- PASS password: used to send the user password to the server.
- LIST: used to ask the server to send back a list of all the files in the current remote directory. The list of files is sent over a (new and non-persistent) data connection rather than the control TCP connection.

Downloaded from CSIT Tutor

- RETR filename: used to retrieve (i.e., get) a file from the current directory of the remote host.
- STOR filename: used to store (i.e., put) a file into the current directory of the remote host.

→ Each command is followed by a reply, sent from server to client.
→ The replies are 3-digit numbers, with an optional message following the number.
→ Similar in structure to the status code and phrase in the status line of the HTTP response message.
→ Some typical relies, along with their possible messages, are as follow :

- 331 username ok, password required
- 125 data connection already open, transfer starting
- 425 can't open data connection
- 452 error writing file.

DNS (Domain Name System)

→ DNS is
  - A distributed database implemented in a hierarchy of DNS servers.
  - An application layer protocol that allows hosts to query the distributed database.

→ DNS protocol runs over UDP and users port 53.
→ DNS is commonly employed by other application-layer protocols including HTTP, SMTP and FTP to translate user supplied hosts names to IP addresses.

Mail Server Aliasing

- Permits Company's mail server and web server to have identical (aliased) hostnames.
- E.g. wlink.com: for mail server and for web server also.
- Load distribution

→ Busy sites are replicated over multiple servers. With each server running on a different end system and each having a different IP address.
→ For replicated web servers, a set of IP address is thus associated with one canonical hostname.

Working Of DNS:

→ In order of the user's host to be able to send on HTTP request message to the web server www.facebook.com , the user's host must first obtain the IP address of www.facebook.com . This is done as follows :

i) The same user machine runs the client side of the DNS application.
ii) The browser extracts the host name www.facebook.com from the URL and passes the host name to the client site of the DNS application.
iii) The DNS client sends a query containing the host name to a DNS server.

Downloaded from CSIT Tutor

iv) The DNS client eventually receives a reply which includes the IP address for the host name.

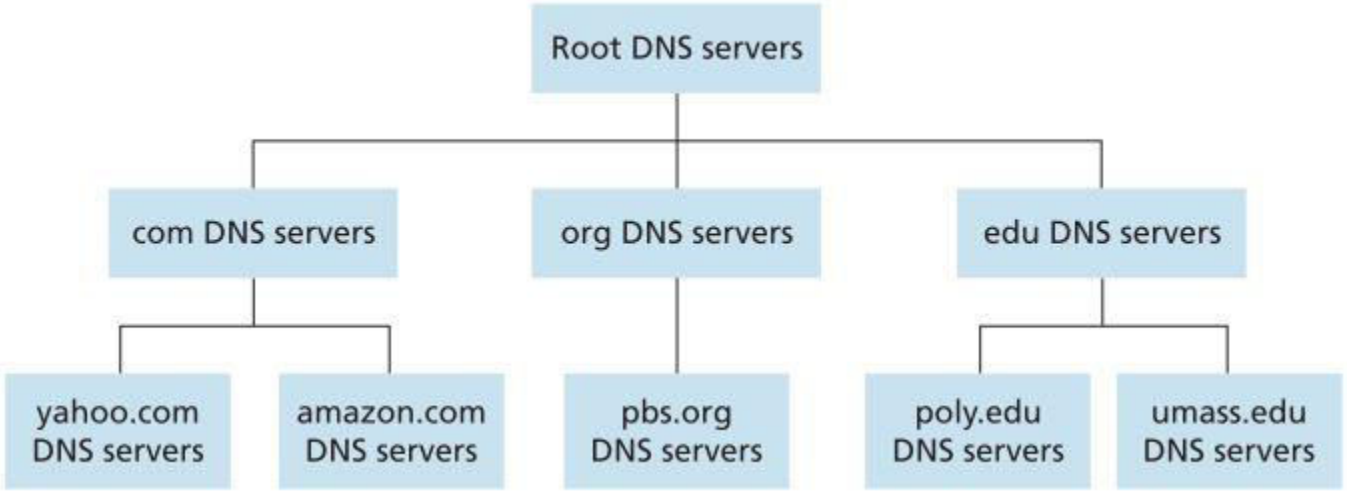
v) Once the browser receives the IP address from DNS, it can initiate a TCP connection to the HTTP server located at port 89 at that IP address. A simple design for DNS world has one DNS server that contains all the mappings. In the centralized design. Clients simply direct all queries to the single DNS server and the DNS server responds directly to the querying clients. The problem with a centralized design include :

- A single point of failure
  If the DNS server crashes, so does the entire internet.
- Traffic volume
  A single DNS server would have to handle all DNS queries.
- Distant centralized database:
  A single DNS server cannot be close to all the querying clients. If we put the single DNS server in US, then all queries from Australia must travel to the other side of the globe, perhaps over slow and congested links. This kind leads to significant delays.
- Maintenance
  The single DNS server would have to keep records for all internet hosts. Not only would this centralized database be huge, but it would have to be updated frequently to account for every new host.

A distributed, Hierarchical database

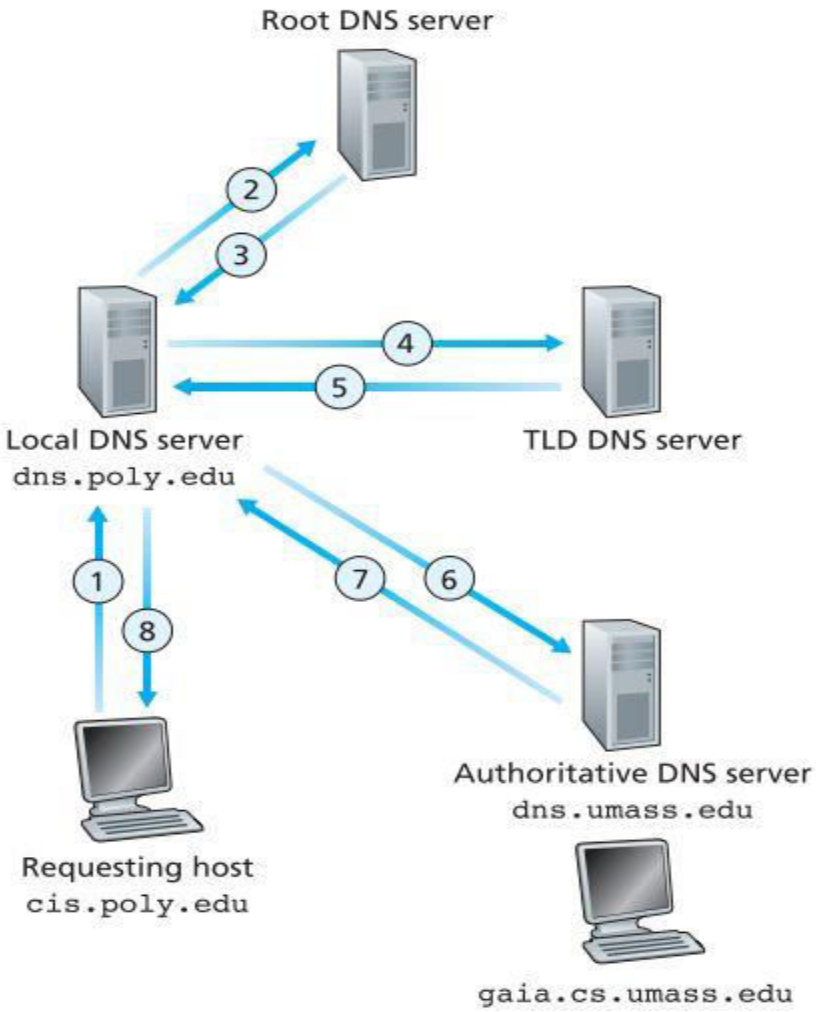
3 classes of DNS servers:

i) Root DNS server
ii) Top-level domain (TLD) DNS servers.
iii) Authoritative DNS servers.



There is another important type of DNS server called the local DNS server.

Downloaded from CSIT Tutor



DNS Records

The DNS servers that together implement the DNS distributed database store resource records (RRs), including RRs that provide hostname to IP address mappings. Each DNS replay message carries one or more resource records.

A resource record is a four tuple that contains the following fields.

(Name, Value, Type, TTL)

TTL is the time to live of the resource record; it determines when a resource should be removed from a cache. In example, we ignore the TTL field. The meaning of Name and Value depend on Type.

There are 5 types of DNS records: **A, CNAME, NS, MX and PTR**

i) Type =A

Address (A) records direct a hostname to a numerical IP address. For e.g., if you want www.urdomain.com to point to IP (which is for e.g. 192.168.0.1) you would enter a record that looks like (www.urdomain.com, 192.168.0.1, A)

Downloaded from CSIT Tutor

ii) Type= CNAME

Canonical name (CNAME) allows a machine to be known by one or more host names. There must be always an A record first, and this is known as canonical name or official name. For e.g.: (www.urdomain.com,192.168.0.1,A)

Using CNAME, you can point other hostnames to the canonical (A record) address. For example:

(fitp.urdomain.com, urdomain.com, CNAME)

(mail.urdomain.com, urdomain.com, CNAME)

(ssh.urdomain.com, urdomain.com, CNAME)

iii) Type=NS

Name server (NS) records specify the authoritative name servers for the domain.

For e.g.: urdomain.com, dns.urdomain.com, NS)

Authoritative server

iv) Type = MX

Mail exchangers (MX) records serve the purpose of using mail server through its web server i.e., canonical name.

For e.g.: (urdomain.com, mail.urdomain.com, MX)

Mail server

v) Type=PTR

Pointer (PTR) records are used for reverse lookups. For e.g.: to make 192.168.0.1 resolve the www.urdomain.com the record would look like (1.0.168.192.in addr.arpa, www.urdomain.com, PTR)

DNS Message (Query and Reply)

Identification	Flags
Number of questions	Number of answer RRs
Number of Authority RRs	Number of Addition RRs
Questions Variable number of questions	
Answers Variable number of Resource Records (RRs)	
Authority	
Additional Information	

Fig: DNS message Format

- The first 12 bytes is the header section which has a number of fields. The first field is a 16-bit number that identifies the query. This identifier is copied into the reply message to a query allowing the client to match received replies with sent queries. There are a number of flags in the flag field. A 1 bit query/reply flag indicates whether the message is a query (0) or a reply (1). A 1-bit authoritative flag is set in a reply message when a DNS server is an authoritative server for a queried name. A 1-bit recursion-desired flag is set when a client (host or DNS server) desires that the DNS server perform recursion when it

Downloaded from CSIT Tutor

doesn't have the record. A 1-bit recursion available field is set in a reply if the DNS server supports recursion. In the header, there are also four numbers of fields. These fields indicate the number of occurrences of the four types of data sections that follow the header.

- The question section contains information about the query that is being made. This section includes
  a) A name field that contains the name that is being queried.
  b) A type field that indicates the type of question being asked about the names for e.g. a host address associated with a name (Type A) or the mail server for the name (Type MX)
- In a reply form a DNS server the answer section contains the resource records for the name that was originally queried. Recall that in each resource record there is the Type (for e.g.: A, NS, CNAME, or MX), the value and the TTL. A reply can return multiple RRs in the answer.
- Since a hostname can have multiple IP addresses (for e.g. for replicated web servers, as discussed earlier in the section). The authority section contains records of other authoritative servers.

The additional section contains other helpful records. For e.g.: the answer field in a reply to an MX query contains a resource record providing the canonical hostname of a mail server. The additional section contains a Type A record providing the IP address for the canonical hostname of the mail server.

Downloaded from CSIT Tutor

1.4 Transport Layer

The primary duties of the transport layer are to transport and regulate the flow of information from a source to a destination, reliably and accurately. End-to-end control and reliability are provided by sliding windows, sequencing numbers and acknowledgements.

To understand reliability and flow control think of someone who studies a foreign language for one year and then visits the country where the language is used. In conversation, words must be repeated for the reliability. People must also speak slowly that the conversation is understood, which relates to flow control.

The transport layer establishes a logical connection between two end points of a network. Protocols in transport layer segment and reassemble data sent by upper layer applications into the same transport layer data string. This transport layer data string provides end-to-end transport services.

Functions:

- Error handling
- Flow control
- Multiplexing
- Connection set-up and release
- Segmentation and reassembly
- Addressing (Port addressing)

Services

- Unreliable unordered unicast or multicast delivery (UDP)
- Reliable, in-order unicast delivery (TCP)

Connectionless Transport: UDP (User Datagram Protocol)

UDP is the connectionless transport protocol in the TCP/IP protocol stack. UDP is a simple protocol that exchanges datagram without guaranteed delivery. It relies on higher layer protocols to handle error and retransmit data.

UDP doesn't use window or Asks reliability is provided by application layer protocols. UDP is designed for applications that do not need to put sequence of segments together.

The following application layer protocols use UDP: TFTP, SNMP, DHCP, and DNS

Hence,

- Used in transport layer
- Offers unreliable connectionless service
- Provides faster service than that of TCP.
- Offers minimum error checking mechanism.
- Supports multicasting because connectionless.
- Offers minimum flow control mechanism.
- Also used by SNMP (Simple Network Management Protocol)

UDP Segment Structure:

Source port number (16)	Destination port number (16)
UDP segment length (16)	UDP checksum (16)
Data	

Downloaded from CSIT Tutor

→ Source port – number of the port that sends data.
→ Destination port – Number of the port that receives data.
→ Length – calculated of bytes in header and data.
→ Checksum – calculated checksum of the header and data field.
→ Data – upper-layer protocol data.

Connection–Oriented Transport: TCP (Transmission Control Protocol)

TCP is a connection-oriented transport layer protocol that provides reliable full-duplex data transmission. TCP is port of the TCP/IP protocol stack. In a connection-oriented environment a connection is established between both ends before the transfer of information can begin. TCP breaks messages into segments, reassembles them at the destination and resends anything that is not received. TCP supplies a virtual circuit between end user applications:
The following application layer protocols use TCP, FTP, HTTP, SMTP and Telnet. Hence
→ This is a real protocol which runs in transport layer.
→ It offers reliable connection-oriented service between source and destination.
→ It acts as if it is connecting two end points together, so that it is a point-to-point connection between two parties.
→ It doesn't support multicasting (because it is connection oriented)
→ The data in TCP is called a segment.
→ Segments are obtained after breaking big files into small pieces.
→ It assists in flow control.
→ It provides buffer to each connection.

TCP segment Structure

<table>
<tr><td colspan="2">Source port (16)</td><td colspan="2">Destination port (16)</td></tr>
<tr><td colspan="4">Sequence number (32)</td></tr>
<tr><td colspan="4">Acknowledgement Number (32)</td></tr>
<tr><td>Header length (4)</td><td>Reserved (4)</td><td>code bits (6)</td><td>window site (16)</td></tr>
<tr><td colspan="2">Checksum (16)</td><td colspan="2">Urgent Pointer (16)</td></tr>
<tr><td colspan="4">Options (0 or 32 if any)</td></tr>
<tr><td colspan="4">Data</td></tr>
</table>

Source port – Number of the port that sends data.
Destination port – Number of the port that receives data.
Sequence number – Number used to ensure the data arrives in the correct order.
Acknowledgement number – next expected TCP octet (defines the number of next byte, a party, expects to receive)
Header length – length of TCP header.
Reserved – reserved for future use.
Code bits – control functions such as setup and termination of a session.

U	A	P	R	S	F

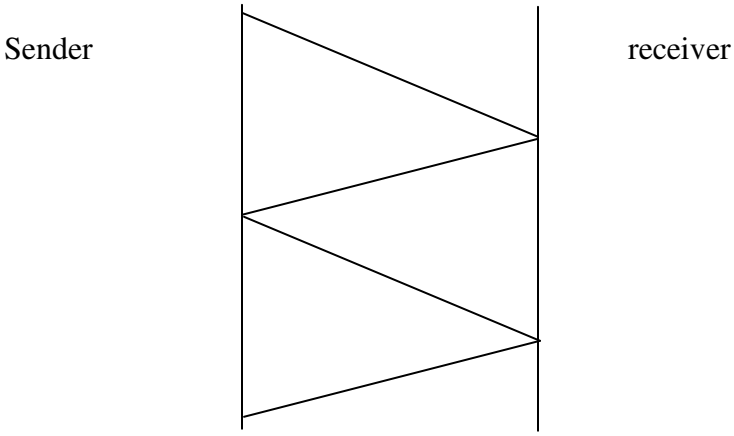
U -urgent valid
A –acknowledges received data

Downloaded from CSIT Tutor

P –data push is valid
R –reset valid
S –synchronization valid (initiates a connection)
F –final valid (terminates a connection)

Window size - number of octets (bytes) that a receiver is willing to accept.
Checksum – indicates the end of the urgent data.
Option – used when sender and receiver negotiate the maximum segment size.
Data – upper-layer protocol data.

Roundtrip Time (RTT) Estimation And Timeout



→ RTT, also called round trip delay is the time required for a signal pulse or packet to travel from a specific source to a specific destination and back again.

→ The sample RRT, denoted as sample RTT, for a segment is the amount of time between when the segment is sent and when an acknowledgement for the segment is received.

→ Obviously, the sample RRT values will fluctuate from segment to segment due to congestion in the routers and to the varying loads on the end systems.

→ In order to estimate a typical RTT, it is natural to take some sort of average of the sample RTT values, called Estimated RRT as

$$\text{Estimated RRT} = (1-x)*\text{estimated RRT} + x*\text{sample RRT}$$

Where x=0.125 recommended.

→ Hence, the new value of estimated RRT is a weighted combination of the previous value of estimated RRT.

→ This weighted average puts more weight on recent samples than the old samples. This is natural, as the more recent samples better reflect the current congestion in the network. In statistics, such an average is called exponential weighted moving average (EWMA).

→ In addition to having an estimate of the RTT, it is valuable to have a measure of the variability of the RTT, called DevRTT an estimate of how much sample RRT typically deviates from estimated RRT.

$$\text{DevRTT} = (1-y)*\text{DevRTT}+y*|\text{sampleRTT}-\text{Estimated RTT}|$$

Where y=0.25 recommended.

Note that, DevRTT is an EWMA of the difference between sample RTT and estimated RTT, if the sample RTT values have little fluctuation, then Dev RTT will be small and vice versa.

Setting and Managing The Retransmission Timeout Interval

Downloaded from CSIT Tutor

Given the values of EstimatedRTT and DevRTT, what value should be used for TCP's timeout interval? Clearly, the interval should be greater than or equal to EstimatedRTT, or unnecessary retransmissions would be sent. But the timeout interval should not be much larger than Estimated RTT; otherwise when a segment is lost, TCP would not quickly retransmit the segment, leading to large data transfer delays. It is therefore desirable to set the timeout equal to the EstimatedRTT plus some margin. The margin should be large when there is a lot of fluctuation in the sample RTT values, it should be small when there is little fluctuation. The value of DevRTT should thus come into play here:

Timeout Interval = estimated RTT + 4 * Dev RTT

Multiplexing and De-multiplexing

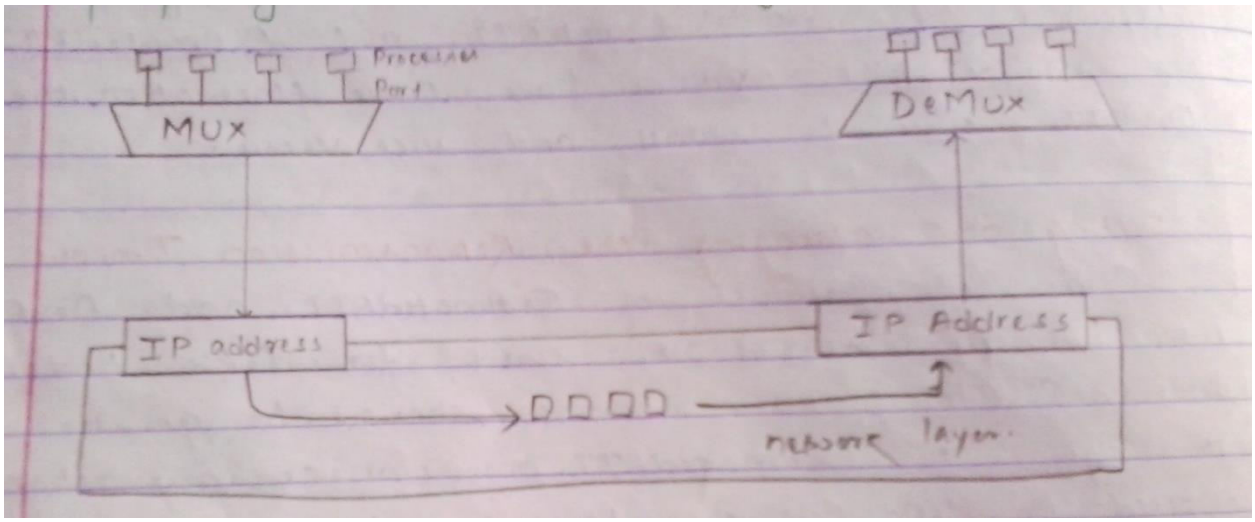


Fig.: Multiplexing and De-multiplexing

→ Extending the host-to-host delivery service provided by the network layer to a process-to-process delivery service application running on the hosts.

→ Consider how a receiving host directs on incoming transport layer segment to the appropriate socket. Each transport layer segments has a set of fields for this purpose. At the receiving end, the transport layer examines these fields to identify the receiving socket and then it directs the segment to that socket. The job of delivering the data in the transport layer segment to the correct socket is called de-multiplexing. The job of gathering data chunks at the source host from the different sockets, encapsulating each data chunk with the header information to create segments and passing the segments to the network layer is called multiplexing.

Flow Control

A TCP connection sets aside a receiver buffer for the connection. When the TCP connection receives bytes that are correct and in sequence, it places the data in the receiver buffer. The associated application process will read data from this buffer, but not necessarily at this instant the data arrives. Indeed, the receiving application may be busy with some other task and may not attempt to read the data until longer after it has arrived. If the application is relatively slow at reading data the sender can very easily overflow the receive buffer by sending too much data quickly.

Downloaded from CSIT Tutor

TCP provides a flow control service to its application to estimate the possibility of the sender overflowing the receive buffer. Flow control is thus a speed-matching service matching the rate at which the sender is sending against the rate at which receiving application is receiving.

Congestion

When too many packets are present in a subnet or a part of subnet, performance degrades. This situation is called congestion. When number of packets dumped into the subnet by the hosts is within its carrying capacity, they are all delivered (except for a few that contain transmission errors), and the number delivered is proportional to the number sent. However, as traffic increases too far, the routers are no longer able to cope, and they begin losing packets. At very high traffics, performance collapses completely and almost no packets are delivered.

Causes of Congestion

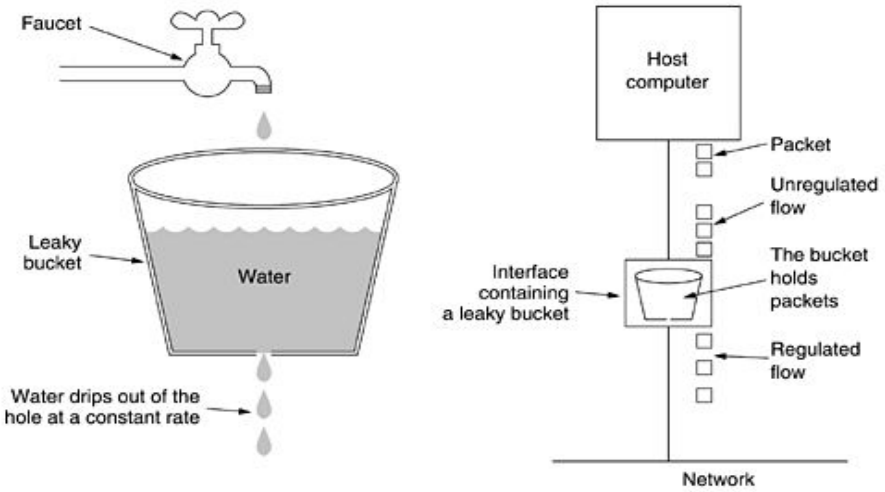
→ When there are more input lines and less or single output lines.
→ When there is slow router i.e., if routers CPU's, are slow
→ If the router has no free buffers i.e., insufficient memory to hold queue of packets.
→ If the components used in subnet (link, router, switches, etc) have different traffics carrying and switching capacities, then congestion occurs.
→ If the bandwidths of the lines are low, it can't carry large volume of packets and caused congestion. Hence, congestion cannot be eradicated but can be controlled.

Congestion Control Algorithms

i) **Leaky Bucket Algorithm**
ii) **Token Bucket Algorithm**
iii) **Choke Bucket Algorithm**

Leaky Bucket Algorithm

Figure (a) A leaky bucket with water. (b) A leaky bucket with packets.



Imagine a bucket with a small hole in the bottom. No matter at what rate water enters the bucket, the outflow is at constant rate, when there is any water in the bucket and zero when the bucket is empty. Also once the bucket is full any additional water entering it spills over the sides and is lost.

Downloaded from CSIT Tutor

The same idea can be applied to the packets conceptually; each host is connected to the network by an interface, containing a leaky bucket, i.e. finite interval queue. If a packet arrives at the queue when it is full, the packet is discarded if one or more processes within the host try to send a packet when a maximum number are already is queue, the new packet is discarded.

Token Bucket Algorithm

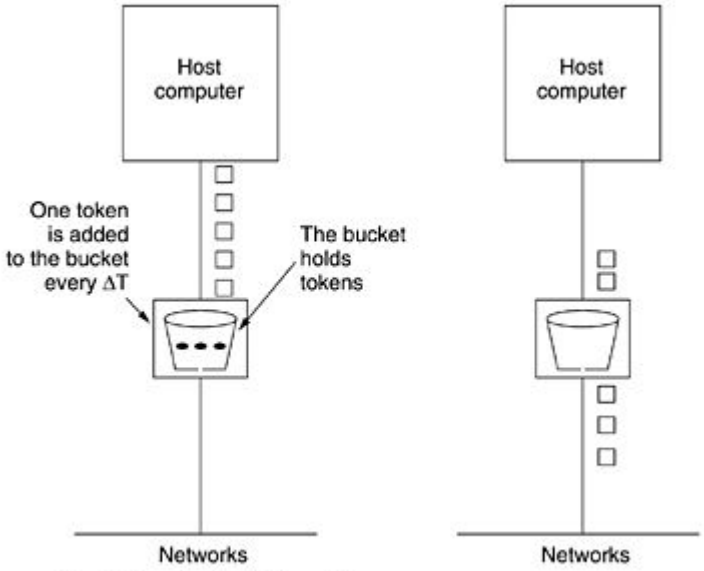


Fig. Token Bucket Algorithm

A leaky bucket algorithm is based on the rigid output pattern at the average rate no matter how bursty the traffic is. In leaky bucket, there are chances of loss of packet as packet is filled in bucket and overflow if bucket is full. To minimize such limitation of bucket, token bucket algorithm was introduced.

The bucket holds token, not packet. Tokens are generated by clock at the rate of one token per ΔT sec. for a packet to be transmitted; it must capture and destroy one token. Token bucket algorithm allows saving up permission to bucket as leaky bucket doesn't allow. This property means that bursts of packets can be sent at once allowing some burstness at output stream and giving faster response to sudden bursts of input.

Another difference between token bucket and leaky bucket is that the token bucket throws away token when the bucket fills up but never discards packets. In token bucket also allows sending bytes basis, for variable size packets. A packet can only be transmitted if enough token are available to cover its length in bytes fractional tokens are kept for further use.

The implementation of basic token bucket algorithm is just a variable counts tokens. The counter is incremented by one at every ΔT sec and decremented by one whenever one packet is sent. When counter hits zero, no packet can be sent.

Reliable Data Transfer (RDT)

RDT is the mechanism where no transferred data bits are corrupted (flipped from o to 1, or vice versa) or lost, and all are delivered in the order in which they were sent.

TCP creates a RDT service on top of IPs unreliable best effort service.

TCP's RDT service ensures that the data stream that a process reads out of its TCP receive buffer is uncorrupted, without gaps, without duplication and in sequence, that is, the byte stream is exactly the same byte stream that was sent by the end system on toe other side of the connection.

Downloaded from CSIT Tutor

Building a RDT protocol

1) **Reliable Data transfer over a Perfectly Reliable Channel**

→ We first consider the simplest case, in which the underlying channel is completely reliable.

→ It is called finite-state machine (FSM).

2) **Reliable Data Transfer over a channel with Bit Errors**

→ A more realistic model of underlying channel is one in which bits in a packet may be corrupted.

→ If receiver receives the packet, the receiver must acknowledge it to the sender whether the packet has received with error-free or not through these:

Positive Acknowledgement (ACK)
Negative Acknowledgement (NAK)

→ If NAK provided, the sender should retransmit the packet.

→ Such protocol is called ARQ (Automatic Repeat Request) protocols.

→ Fundamentally, three additional protocol capabilities are required in ARQ protocols to handle the presence of bit errors :

- **Error Detection**

→ Internet checksum field

→ Error-detection and correction techniques

→ Require extra bits (beyond the bits of original data to be transferred) to be sent from the sender to the receiver, these bits will be gathered into the packet checksum field.

- **Receiver Feed Back**

→ Receiver provides feed back

→ Positive (ACK) -1 value

→ Negative (NAK) – 0 values

- **Retransmission**

→ A packet that is received in error at the receiver will be retransmitted by the sender.

These phenomena are called **stop-and-wait protocols**.

An amazing case will occur if ACK or NAK is corrupted i.e. the sender could not get the feedback from sender.

Consider three probabilities for handling corrupted ACKs or NAKs.

- A second alternative is to add enough checksum bits to allow the sender not only to detect, but also to receiver from bit errors. This solves immediate problem for a channel that can corrupt packets but not lose them.
- A third approach is for the sender simply to resend the current data packet when it receives a garbled ACK or NAK packet. This introduces duplicate packets into the sender-to-receiver channel. The receiver doesn't know whether the ACK or NAK it last sent was received correctly at the ender. Thus, it cannot know whether an arriving packet contains new data or is a retransmission.

→ A solution to this problem is to a new field called "sequence number" to the data packet.

→ For this stop-and-wait protocol, a 1-bit sequence number will be ok.

3) **Reliable Data Transfer Over A lossy Channel with Bit Errors**

Downloaded from CSIT Tutor

→ Suppose now that in addition to corrupting bits, the underlying channel can lose packets as well.
→ The sender must get information of packet loss on the way from the receiver so that the sender can retransmit.
→ The sender must clearly wait at least as long as a round-trip delay between the sender and the receiver.
→ If ACK is not received within this time, the packet is retransmitted.
→ If a packet experiences a particularly large delay, the sender may retransmit the packet even though neither the data packet nor its ACK have been lost.
→ This introduces the possibility of duplicate data packets in the sender-to-receiver channel.
→ For all this, we can do is retransmit.
→ But implementing a time-based retransmission mechanism requires a "countdown timer" that an interrupt the sender after a given amount of time has expired.
→ The sender will thus need to be able to
- Start the timer each time a packet (either a first time packet or a retransmission) is sent.
- Respond to a timer interrupt (taking appropriate actions)
- Stop the timer.

Checksums: sequence numbers, timers, positive and negative acknowledgement 1 (page 254)

Reliable Data Transfer Protocol

1. **Pipelined**
2. **Go-Back-N(GBN)**
3. **Selective Repeat (SR)**

Pipelined Reliable Data Transfer Protocol:

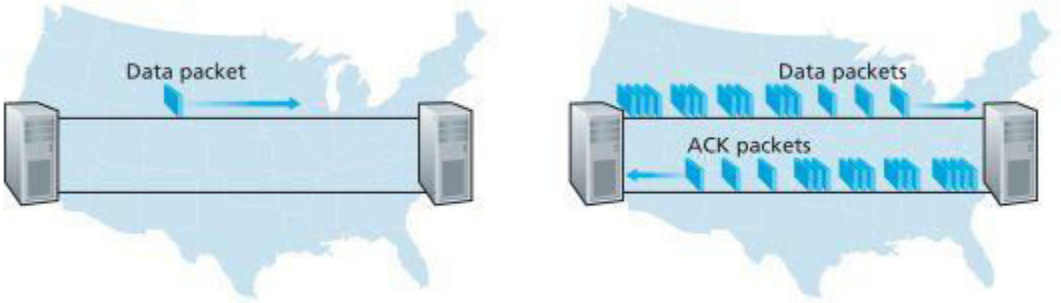


Fig: stop-and-wait vs. pipelined protocol

Instead of sending a single packet in stop and wait manner, the sender is allowed to send multiple packet without waiting for acknowledgements, as illustrate in fib (b). fig (b) shows that if the sender is allowed to transmit three packets before having to wait for acknowledgement the utilization of the sender is essentially tripled. Since the many in-transmit sender-to-receiver packets can be visualized as a filling a pipeline, this technique is known as pipelining.

Consequences of pipelined protocol

- Increment in the range of sequence numbers.

Downloaded from CSIT Tutor

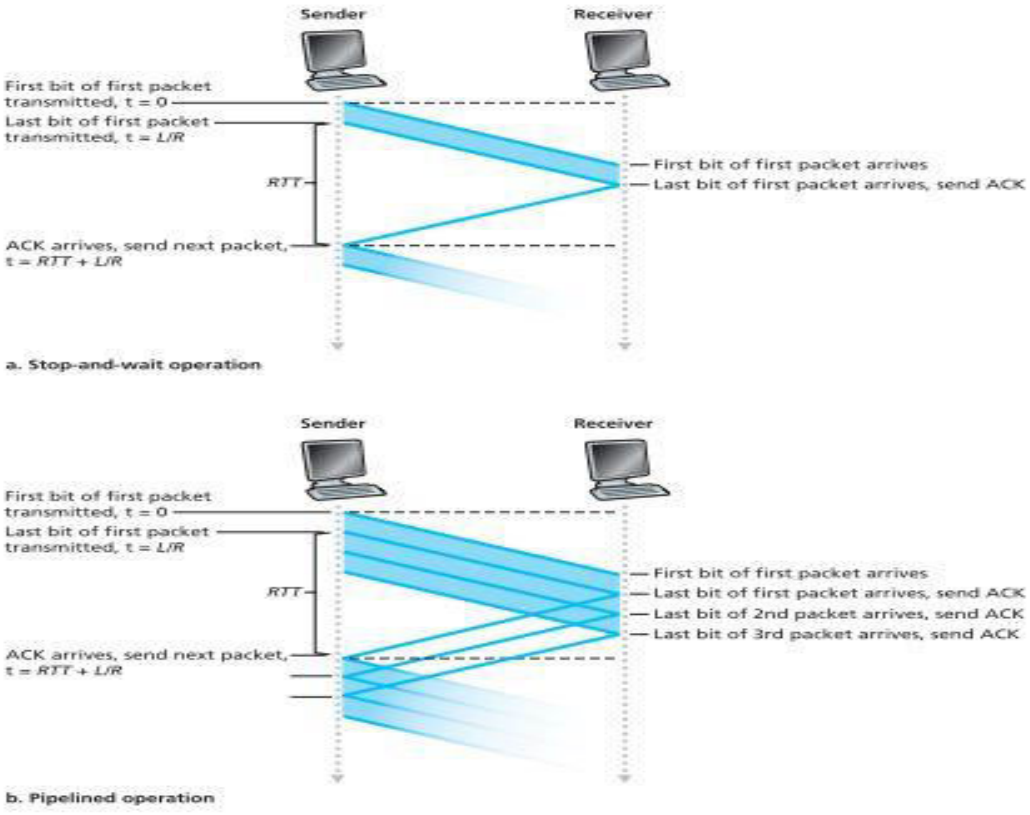
- Sender and receiver have to buffer more than one packet.
- Range of sequence numbers and the buffering requirements will depend on the manner in which a data transfer protocol responds to lost, corrupted and overly delayed packets.

Go-Back-N (GBN):

In a GBN, protocol, the sender is allowed to transmit multiple packets (when available) without waiting for an acknowledgement but is allowed to have no more than some maximum allowable number, N, of an un acknowledged packets in the pipeline.

Figure shows the sender view of range of sequence numbers in a GBN protocol. If we define base to be the sequence number of the oldest un acknowledge packet and next sequence num to be the smallest unused sequence numbers (i.e. the sequence number of the next packet to be sent), then four intervals in the range of sequence numbers can be identified. Sequence numbers in the interval [o, base-1] correspond to packets that have already been transmitted and acknowledged. The interval [base, next sequence-1] corresponds to packets that have been sent but not yet acknowledged. Sequence numbers in the interval [next sequence, base+N-1] can be used for packets that can be sent immediately, should data arrive from the upper layer. Finally, sequence number greater than or equal to base+ N cannot be used until an unacknowledged packet currently in the pipeline (specifically, the packet with sequence number base) has been acknowledged.

As suggested by figure, the range of permissible sequence numbers for transmitted but not yet acknowledged packets can be viewed as a window of size N over the range of sequence numbers. As the protocol operates, this window slides forward over the sequence number space. For this reason, N is often referred to as the window size and the GBN protocol itself as a sliding window protocol.



Downloaded from CSIT Tutor

Selective Repeat (SR):

GBN itself suffers from performance problems. Many packets can be in the pipeline when the window size and bandwidth-delay product are both large. A single packet error can thus cause GBN to retransmit a large number of packets, many unnecessarily. As the probability of channel error increased, the pipeline can become filled with these unnecessary retransmissions.

As the name suggests, selective- repeat protocols avoid unnecessary retransmissions by having the sender retransmit only those packets that it suspects were received in error (i.e., were lost or corrupted) at the receiver. A window size of N will again be used to limit, the number of outstanding, unacknowledged packets in the pipeline. However, unlike GBN, the sender will have already received ACKs for some of the packets in the window.

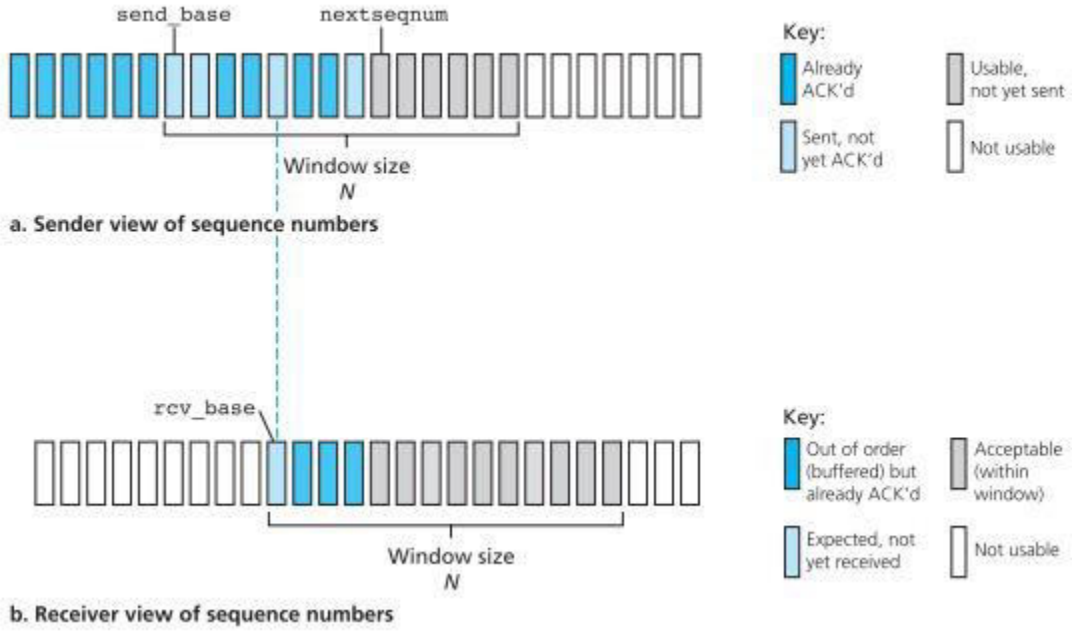
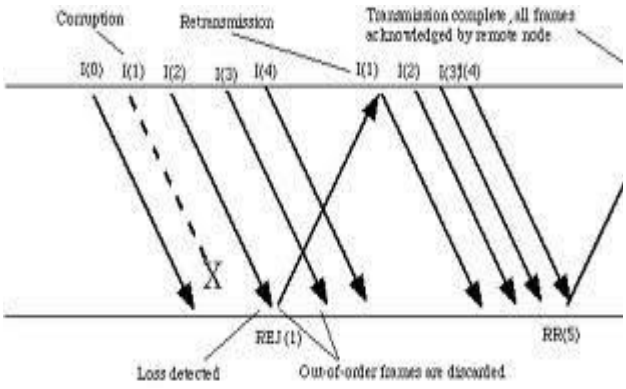
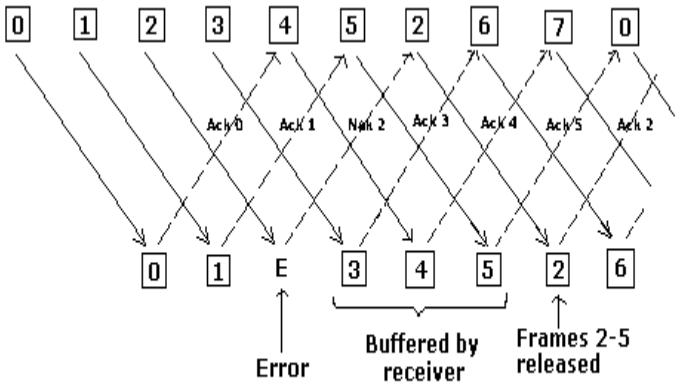


Fig.: Selective-repeat (SR) sender and receiver views of sequence-number space

The SR receiver will acknowledge a correctly received packet whether or not it is in order. Out of order packets are buffered until any missing packets (i.e. packets with lower sequence numbers) are received at which points a batch of packets can be delivered in order to the upper layer.





Downloaded from CSIT Tutor

Downloaded from CSIT Tutor

1.5 Network layer (Internet layer)

Functions

- **Path determination:** route taken by packets from source to destination (Routing Algorithm).
- **Forwarding:** more packets from router's input to appropriate router output.
- **Call setup:** some n/w architectures require router cell setup along the path before data flows.

The following protocol operate at the TCP/IP internet layer

i. **Internet protocol (IP):** IP provides connectionless, best- effort delivery routing of packet. IP is not concerned with the contents of the packets but looks for a path to the destination.
ii. **Internet control message protocol (ICMP):** ICMP Provides control and messaging capabilities.
iii. **Address Resolution Protocol (ARP):** ARP determines the data link layer address or MAC address, for known IP address.
iv. **Reverse ARP (RARP):** RARP determines the IP address for known MAC address.

Network service model

It means the characteristics of end to end transport of packets between sending and receiving end system. In the sending host, when the transport layer passes a packet to the network layer, specific services that could be provided by the network layer include:

- Guaranteed delivery
- Guaranteed delivery with bounded delay.

Furthermore, the following service could be provided to a flow of packets between a given source and destination:

- In order packet delivery
- Guaranteed minimal bandwidth
- Guaranteed maximum jitter
- Security service.

Virtual Circuit and Datagram Networks

The internet transport layer provides each application a choice between two services UDP (a connectionless service) or TCP (a connection oriented service). In similar manner, a network layer can also provide connectionless service (datagram networks) or connection service (virtual circuit network)

Although these transport layer and network layer service models seem parallel, there are some crucial differences:

i. In transport layer, it is process to process service. But, in network layer, it host to host service.

Downloaded from CSIT Tutor

ii. In all computer network architectures up to now (internet, ATM, frame relay, and soon), the network layer provides either a host to host connection service or host to host connectionless service but not both.

iii. Connection oriented service in transport layer in implemented at the edge of the network in the end systems; however, the network layer connection service is implemented in the network core as well as the end system.

Virtual Circuit (VC) Network

Many network architectures (not internet) including those of ATM and frame relay are VC network and therefore, use connections at the network layer. These network layer connections are called virtual circuits (VCs). Let's now consider how a VC service can be implemented in a computer network.

A VC consists of

1) A path (i.e. a series of links and routers) between the source and destination hosts.
2) VC numbers, one number for each link along the path.
3) Entries in the forwarding table in each router along the path.

A packet belonging to a virtual circuit will carry a VC number in its header. Because a virtual circuit may have a different VC number on each link, each intervening router must replace the VC number of each traversing packet with a new VC number. The new VC number is obtained from the forwarding table.

There are three identifiable phases in a virtual circuit

i) VC setup
ii) Data transfer
iii) VC teardown

Datagram network

Internet is a datagram network in which each time an end system wants to send a packet, it stamps that packet with the address of the destination end system and then pops packet into the network. Routers in a datagram network don't maintain any state information about VCs.

As a packet is transmitted from source to destination, it passes through a series of routers. Each of these routers uses the packets destination address to forward the packet. Specifically, each routers has a forwarding table that maps destination addresses to link interfaces, when a packet arrives at the router, the router uses the packets destination address to look up the appropriate output link interface in the forwarding table. The router then intentionally forwards the packet to that output link interface.

Routing

Once you create an inter network by connecting your WANs and LANs to a router. You'll need to configure local network addresses, such as IP addresses, to all hosts on the internet work so that they can communicate across that internetwork.

Downloaded from CSIT Tutor

The term routing refers to taking a packet from one device and sending it through the network to another device on a different network. Routers don't really care about hosts. They only care about networks and the best path to each network. The logical network address of the destination host is used to get packets to a network through a routed network, and then hardware address of the host is used to deliver the packet from a router to the correct destination host.

Principles: If your network has ho routers, then it is clear that you are not routing. Routers route traffic to the entire network in your internetwork. To be able to route packets, a router must know, at minimum, the following:

→ Destination address
→ Neighbor routers from which it can learn about remote network.
→ Possible routers to all remote networks.
→ The best route to each remote network.
→ How to maintain and verify routing information.

The router learns about remote network from neighboring routers or from an administrator. The router then builds a routing table (a map of the internet work) that describes how to find the remote network. If the network is directly connected, the router already knows how to get to it.

Static Vs Dynamic Routing

If a network is not directly connected to the router the router must use one of two ways to learn how to get to the remote network: static routing or dynamic routing.

Static routing means someone must hand-type all network locations into the routing table. If static routing is used, the administrator is responsible for updating all changes by hand onto all routers.

In dynamic routing, a protocol acts on all neighboring routers. Then the routers update each other about all the networks they know about and place this information into the routing table. If a change occurs in the network, the dynamic routing protocols automatically inform all routers about the event e.g. RIP V1, RIP v2, OSPF, EIGRP.

Routing algorithm: Distance vector vs. link state

There are three classes of routing protocols:

i) Distance vector

The distance-vector protocols are in use today. Find the best path to a remote network by judging distance. For e.g., in the case of RIP routing, each time a packet goes through a router, that's called a hop. The route with the least number of hops to the network is determined to be the best route. The vector indicates the direction to the remote network. E.g.: RIP, IGRP, they periodically send the entire routing table to directly connected neighbors.

ii) Link State

It is also called shortest-path-first protocols in which the routers each create three separate tables. One to keep track of directly attached neighbors, one determines the topology of

Downloaded from CSIT Tutor

the entire internet work, and one is used as the routing table. Link-state routers know more about the internet work than any distance-vector routing protocol. E.g. OSPF (Open Shortest Path First). They send updates containing the state of their own links to all other directly connected routers on the network, which is then propagated to their neighbors.

iii) Hybrid

Hybrid protocols use aspects of both distance vector and link state. E.g.: EIGRP.

Hierarchical Routing: intra-AS routing and inter-AS routing

Autonomous system (AS) is a collection of networks under a common administrative domain, which basically means that all routers sharing the same routing table information are in the same AS

According to AS, there are two types of routing protocols:

i. Intra-AS routing/interior Gateway protocol (IGP) e.g.: RIP, OSPF.
ii. Inter-AS routing/exterior Gateway protocol (EGP) e.g.: Border gateway protocol (BGP)

The internet Protocol (IP)

IP is sometimes referred to as an unreliable protocol. This does not mean that IP will not accurately deliver data across a network. IP is unreliable because it does not perform error checking and correction. That function is handled by upper layer protocols from the transport or application layers.

IP performs the following operations:

→ Defines a packet and an addressing scheme
→ Transfers data between the internet layer and network access layer.
→ Routers packets to remote hosts.
→ The main function of IP is forwarding and addressing in the internet.

IPv4 Addressing

A router's job is to receive a datagram on one link and forward the datagram on some other link, a router necessarily has two or more links to which it is connected. The boundary between the router and any one of its link is called an interface. Because every host and router is capable of sending and receiving IP datagram, IP requires each host and router interface to have its own. IP address thus, an IP address is technically associated with an interface, rather than with the host router containing that interface.

Each IP address is 32 bits long (4 bytes) and thus a total of 2^32 possible IP address. Approximately, there are about 4 billion possible IP addresses. These IP addresses are typically written in so called dotted-decimal notation, in which each byte of the address is written in its decimal from and is separated by a period (dot) from other bytes in the address.

For e.g. : consider the IP address 192.168.10.5 the 192 is the decimal equivalent of the first 8 bits of the address, so are the 168, 10 and 5. Thus, the address 192.168.10.5 in binary notation is

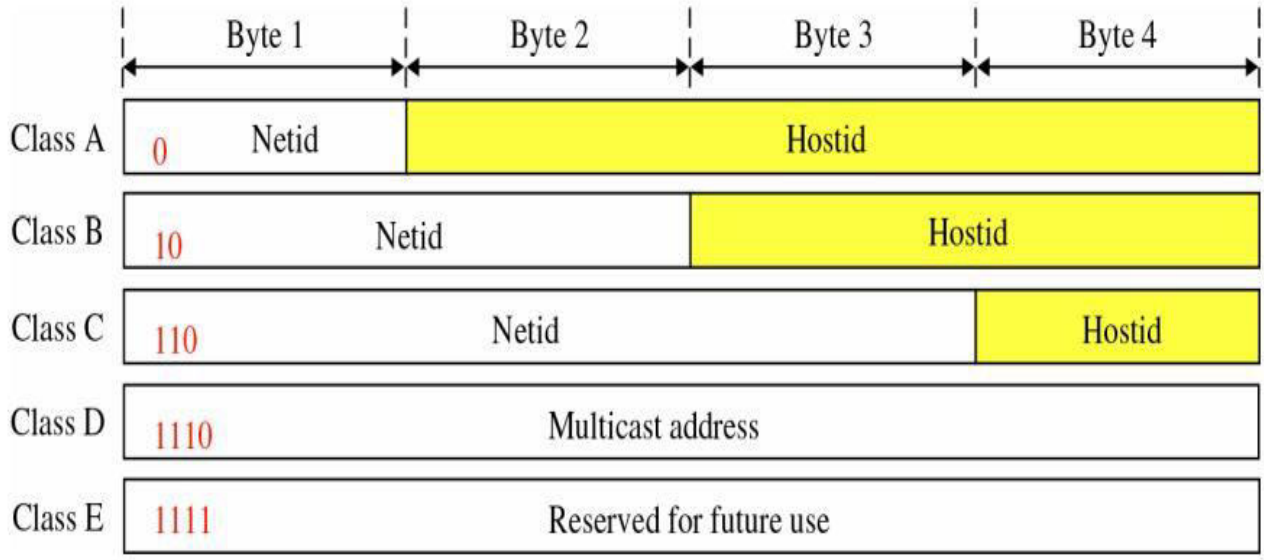
11000000 101010000 00001010 00000101

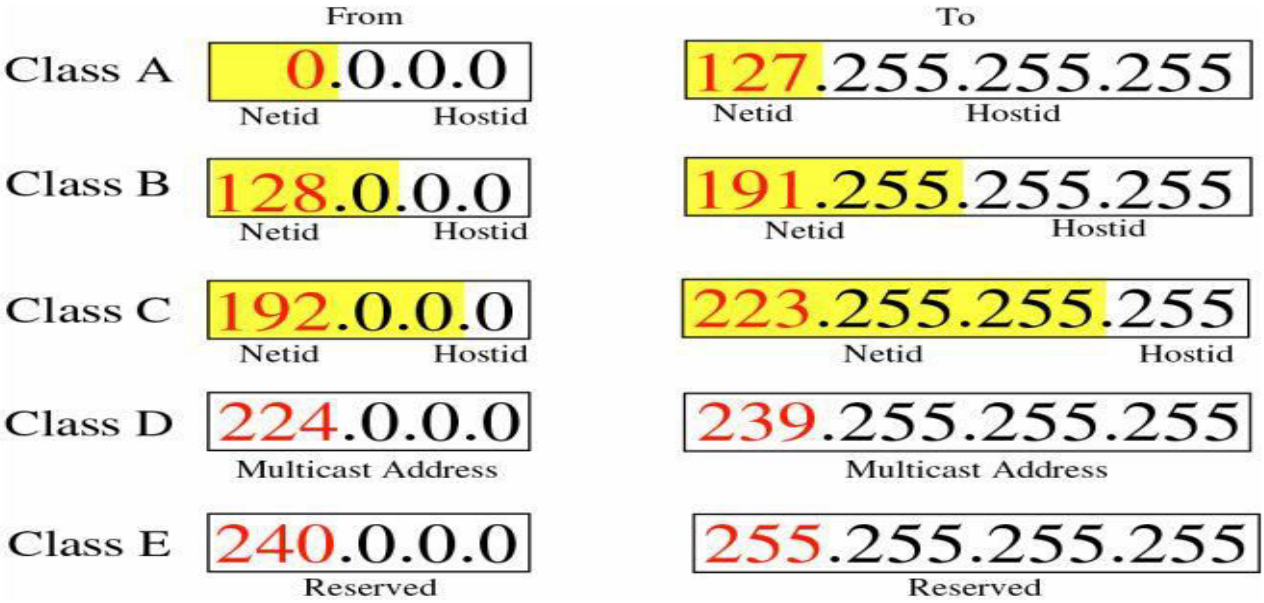
Downloaded from CSIT Tutor

Each interface an every host and router in the global internet must have an IP address that is globally unique (except for interfaces behind NATs). A portion of an interface's IP address will be determined by the subnet to which it is connected.

Different classes of IPV4 address

An internet address is made of 4 bytes (32 bits) that define a host's connection to a network.
IP address is made up of (netid + hostid)





Class A

- Range: 0 – 127
- So total of 126 (2^{8-1}) Networks are possible and total host = 2^{24} in each Network.

Downloaded from CSIT Tutor

- Default subnet mask is 255.0.0.0

Class B
- Range: 128 – 191
- So total of 2^{16-2} Networks are possible and total host = 2^{16} in each Network.
- Default subnet mask is 255.255.0.0

Class C
- Range: 192 – 223
- So total of 2^{24-3} Networks are possible and total host = 2^{8} in each Network.
- Default subnet mask is 255.255.255.0

Class D
- Range: 224 – 239
- Used for Multicasting
- E.g. 224.0.0.1 (group)

Class E
- Range 240-255
- Not used (for future use)

Private Vs Public Address

The people who created the IP addressing scheme also created the IP addressing scheme also created what we call private IP addresses which can be used on a private network, but they are not routable through the Internet. This is designed for the purpose of creating a measure of well-needed security, but it also conveniently saves valuable IP address space.

To accomplish the connection between the ISP and the corporation, the end user, no matter who they are need to use something called Network Address Translation (NAT), which basically takes a private IP address and converts it use on the internet. Many people can use the some real IP address to transmit out onto the internet. Doing things this way saves megatons of address space-good for us all. The reserved private addresses

Class A: 10.0.0.0 through 10.255.255.255

Class B: 172.16.0.0 through 172.16.255.255

Class C: 192.168.0.0 through 192.168.255.255

IP Datagram Format

Different s field used in IP (Version 4) datagram are depicted in fig below:

<table>
<tr><th>Version (4)</th><th>HLEN (4)</th><th>Types of services (8)</th><th colspan="2">Datagram Length (16)</th></tr>
<tr><td colspan="3">Identifier (16)</td><td>Flags (3)</td><td>Fragment Offset (13)</td></tr>
<tr><td colspan="2">TTL (8)</td><td>Protocol (8)</td><td colspan="2">Header Checksum (16)</td></tr>
<tr><td colspan="5">Source IP address (32)</td></tr>
<tr><td colspan="5">Destination IP address (32)</td></tr>
<tr><td colspan="5">Options or Padding not always</td></tr>
<tr><td colspan="5">Data (variable)</td></tr>
</table>

Downloaded from CSIT Tutor

* Number in bracket indicates bits used in that field.

Version: Identifies the version of IP in use. Current version is IPV4.
HLEN: Header length is set to a value to indicate the length of datagram header. Most IP datagram doesn't contain options, so HLEN mostly indicates where the data begins in datagram. Typical IP datagram has 20 bytes header.
Types of services: Identifies different types of services included in IP datagram such as delay, throughput, precedence etc. IP datagram can be real-time or non-real-time as per type of services
Datagram Length: Indicates total length (Data + Header) of the IP datagram. Maximum length if IP datagram is 2^{16}=65535 bytes but in general not more than 1500 bytes.
Identifiers / Flags / Fragment Offset: Identifier (also called Fragment ID) indicates all fragments that belong together. Flags indicate that other fragments to follow. All fragments except last are indicated as 1 and last flag is 0.Fragment offset is used to tell the receiving host how to reassemble the packets.
Time-to-Live (TTL): TTL is used to measure the time a datagram has been in internet. Each Gateway in internet checks this field and discards packet if TTL is 0.
Protocol: this field is used to indicate upper layer protocols (Transport layer) that are to receive the datagram at the destination host. Either TCP or UDP receive the IP datagram at destination.
Header Checksum: Used o detect bit error at the receiving datagram.
Source/Destination address: IP datagram used two 32-bits addresses called source IP address and Destination IP address.
Options: The option field is not used in every datagram. His field is used sometimes for network management and diagnostics.
Data: Data field contains the user data. IP stipulates that the combination of header and Data can't exceed 65535 bytes. Data length varies from protocol to protocol used in network access layer.

IP datagram Fragmentation

Not all network access layer protocols can carry packets of the same size. Some protocols can carry big packets and other protocols can carry small packets. For example, Ethernet packets can carry no more than 1500 bytes of data, where as packets for many wide area network are not more than 576 bytes. The maximum amount of the data that the network access layer (TCP/IP model) protocol can carry is called Maximum Transfer Unit (MTU). Because each IP datagram is encapsulated within the network access layer packet for transport between routers, the MTU of the network access protocol places a hard limit on the size of an IP datagram. The main problems here are that each of the links along the route between sender and receiver can use different network access protocols, and each of these protocols can have different MTUs.

When the size of IP datagram is large than the MTU of Network access layer protocols, this IP datagram need to be fragmented into two or more fragments. These fragments need to be

Downloaded from CSIT Tutor

reassembled before they reach to the destination transport layer. Reassembling is done with fragment ID and Fragment Offset. Indeed, both TCP and UDP are expecting to receive complete unfragmented segments from the Internet layer.

The designers of IPV4 felt that the fragmenting, reassembling and possibly again fragmenting and reassembling datagram into the routers would introduce significant complication into the protocol and put a damper on router performance. Fragmentation and reassembly add extra burden at sending routers and receiving hosts. So fragmentation should be minimized as far as possible. This is often done by limiting the TCP /UDP segments to a relatively small size i.e. less than 576 bytes (all network access layer protocols supported by IP are supposed to have MTUs at least 576 bytes. Fragmentation can be entirely eliminated by using an MSS (maximum segment size) of 536 bytes, 20 bytes for TCP header and 20 bytes for IP header.

Fragmentation is supported by only IPV4 not by IPV6.

Features of IP:

- *It is connectionless service:* So without prior call setup, it permits to exchange traffics between two host computers.
- *Datagram could be lost*: As IP is connectionless; it is possible that datagrams could be lost between two end user's stations.
- *IP hides underlying sub network from the end user:* In this context, it creates a virtual network for the end user. This aspect of IP is quite attractive, because it allows different types of networks to attach to an IP gateway. As a reason IP is reasonably simple to install and, because of its connectionless design, it is quite accommodating.
- *IP is unreliable, best effort and datagram type protocol:* It has no reliability mechanisms. It has no error recovery procedures for the underlying sub networks.
- *IP has no flow control mechanisms:* The user datagram may lost , duplicated or even arrive at out of order. It is not the job of IP to deal with most of these problems. It is not the job of IP to deal with most of these problems, as most of the problems are passed to the next upper layer, TCP.
- *IPV4 supports fragmentation:* Fragmentation refers to an operation where in a protocol data unit (PDU) is divided or segmented into smaller units.

Subnetting

→ A subnetwork, or subnet, is a logically visible subdivision of an IP network. The practice of dividing a network into two or more networks is called subnetting.

→ All computers rhat belong to a subnet are addresses with a common, identical, most significant bit group in their IP address. This results in the logical division of an IP address into two fields,

- A network or routing prefix
- The rest field or host identifier

→ The rest field is an identifier for specific host or network interface.

Downloaded from CSIT Tutor

Address class	Bits for subnet mask	Network prefix
A	11111111 00000000 00000000 00000000	/8
B	11111111 11111111 00000000 00000000	/16
C	11111111 11111111 11111111 00000000	/24

Benefits of subnetting

→ Reduced network traffic

→ Simplified management

→ Smaller broadcast domain

Subnet mask

A subnet mask is a 32-bit number that masks an IP address, and divides an IP address into network address and host address. Subnet mask is made by setting the network bits to all 1's and setting host bit to all 0's. Within a given network, two host addresses are reserved for special purpose. The '0' address is assigned a network address and '255' is assigned to a broadcast address, and they cannot be assigned to hosts.

Network address – Used to identify the network itself .Data that is sent to any host on that network (198.150.11.1- 198.150.11.254) will be seen outside of the local area network as 198.159.11.0. The only time that the host numbers matter is when the data is on the local area network.

Broadcast address – Used for broadcasting packets to all the devices on a network. Data that is sent to the broadcast address will be read by all hosts on that network. The Broadcast Address for above IP addresses is 198.150.12.255.

CIDR (Classless Inter Domain Routing)

CIDR was introduced in 1993 replacing the previous generation of IP address syntax – classful networks. CIDR allowed for more efficient use of IPv4 address space and prefix aggregation, known as route summarization or supernetting.

CIDR allows routers to group routes together to reduce the bulk of routing information carried by core routers. With CIDR, IP addresses and their subnet mask are written as four octets, separated by periods, followed by a forward slash (/) and a two digit number that represents the network mask.

e.g. 10.1.1.0/30

172.16.1.16/28

192.168.1.32/27

Downloaded from CSIT Tutor

ICMP (Internet Control Message Protocol)

The internet protocol is connectionless-mode protocol, and as such, it has no error reporting and error-correcting mechanisms. It relies on a module called the Internet control message protocol (ICMP) to;

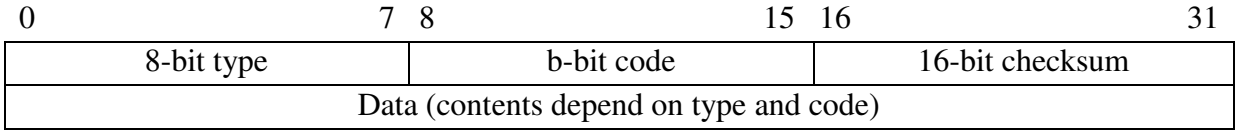
a. Reports errors on the processing of a datagram
b. Provide for some administrative and status messages.

ICMP sends messages and reports errors to the source host regarding the delivery of a packet. ICMP notifies the host if a destination is unreachable. ICMP is also responsible for managing and creating a time-exceeded message in the event that the lifetime of the datagram expires. ICMP also performs certain editing functions to determine if the IP header is in error or otherwise unintelligible.

The error and status reporting services of ICMP are summarized as below.

Type	Code	description
0	0	echo reply (ping)
3	0	destination network unreachable
3	1	destination host unreachable
3	2	destination protocol unreachable
3	3	destination port unreachable
3	6	destination network unknown
3	7	destination host unknown
4	0	source quench (congestion control - not used)
8	0	echo request (ping)
9	0	route advertisement
10	0	router discovery
11	0	TTL expired
12	0	bad IP header

ICMP packet format



Type: type of message
Code: Subtype of message
Checksum: 1's complement computed over entire ICMP message (except for the checksum field itself, which is set to zero)
Data: depends on type and code

NAT (Network Address Translation)

NAT (Network Address Translation or Network Address Translator) is the translation of an Internet Protocol address (IP address) used within one network to a different IP address known within another network. One network is designated the *inside* network and the other is the *outside*. Typically, a company maps its local inside network addresses to one or more global

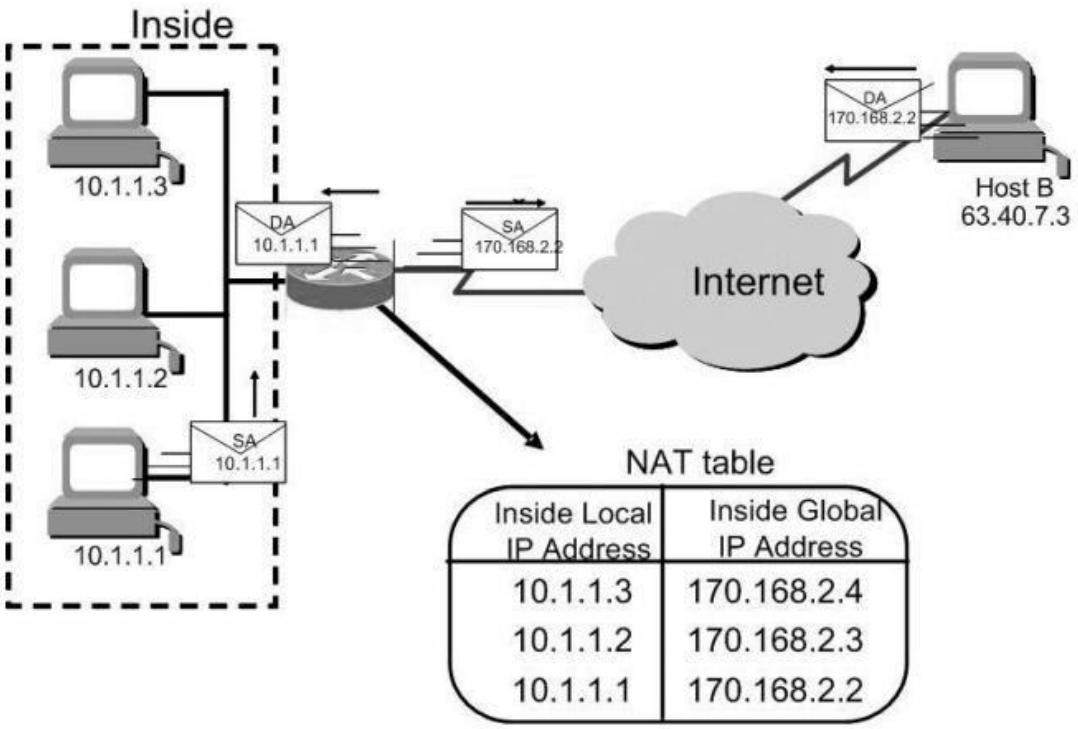
Downloaded from CSIT Tutor

outside IP addresses and unmaps the global IP addresses on incoming packets back into local IP addresses. This helps ensure security since each outgoing or incoming request must go through a translation process that also offers the opportunity to qualify or authenticate the request or match it to a previous request. NAT also conserves on the number of global IP addresses that a company needs and it lets the company use a single IP address in its communication with the world.

NAT is included as part of a router and is often part of a corporate firewall. Network administrators create a NAT table that does the global-to-local and local-to-global IP address mapping. NAT can also be used in conjunction with *policy routing*. NAT can be statically defined or it can be set up to dynamically translate from and to a pool of IP addresses.

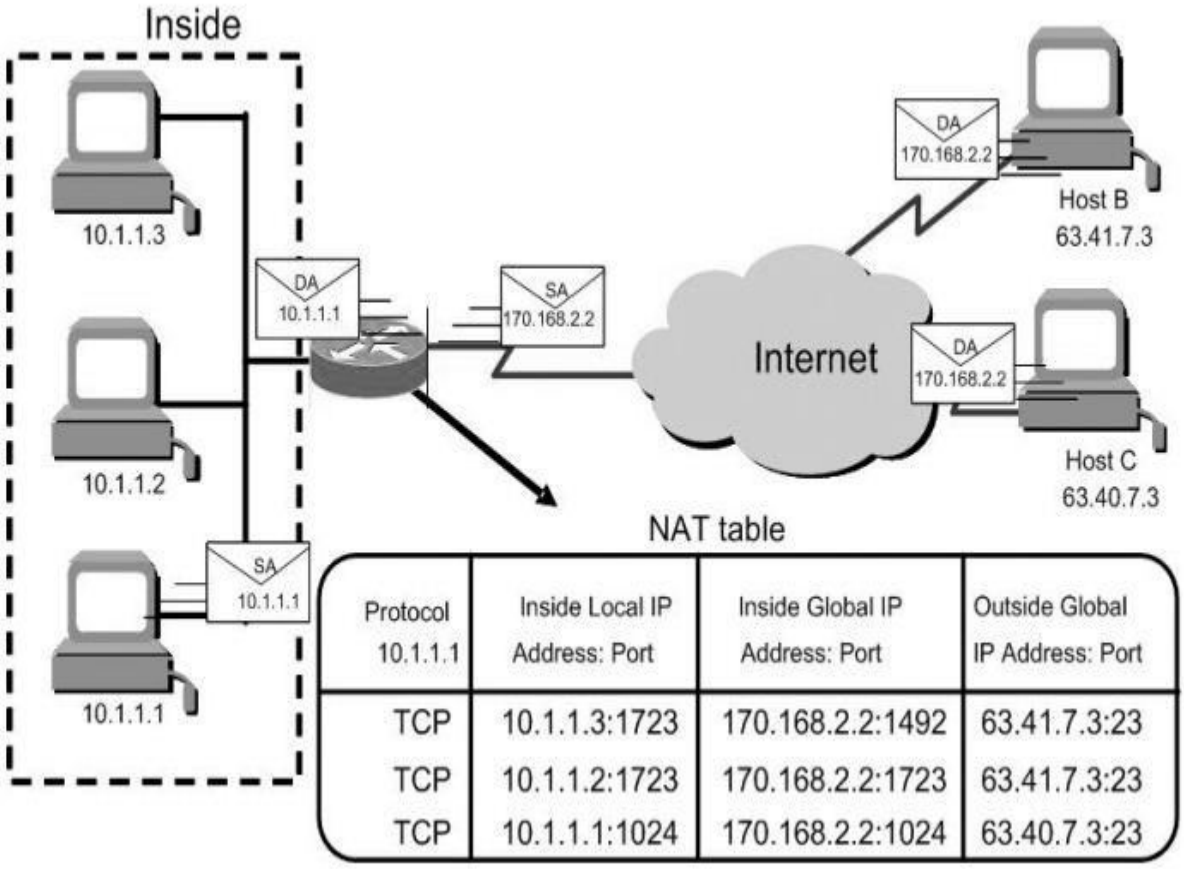
Types of NAT

- **Static NAT**: A local IP address to one global IP address statically



- **Dynamic NAT:** A local IP address to any of a rotating pool of global IP addresses that a company may have

Downloaded from CSIT Tutor



- **NAT Overloading (PAT – Port Address Translation):** A local IP address plus a particular TCP port to a global IP address or one in a pool of them.

NAT Terms

- **Inside local address**—Name of inside source inside translation
- **Outside local address**—Name of destination host before translation
- **Inside global address**—Name of inside host after translation
- **Outside global address**— Name of outside destination host after translation

Need of NAT

i) You need to connect to the internet and your hosts don't have globally unique IP addresses.

ii) You change to a new ISP that requires you to renumber your network.

iii) You need to merge two intranets with duplicate addresses.

Routing in the Internet

1. RIP (Routing Information Protocol)

→ Distance-vector routing protocol.

Downloaded from CSIT Tutor

→ Intra AS routing Protocol.
→ Employs the hop count as a routing metric.
→ RIP prevents routing loops by implementing a limit on the number of hops allowed in a path from the source to a destination.
→ The maximum number of hops allowed for RIP is 15.
→ This hop limit, however, also limits the size of networks that RIP can support.
→ A hop count of 16 is considered an infinite distance and used to deprecate inaccessible, inoperable, or otherwise undesirable routes in the selection process.
→ Periodic updates every 30 seconds, even the topology not changed.
→ In the early deployments, routing tables were small enough that the traffic was not significant. As networks grew in size, however, it became evident there could be a massive traffic burst every 30 seconds, even if the routers had been initialized at random times
→ RIP uses the User Datagram Protocol (UDP) as its transport protocol, and is assigned the reserved port number 520
→ Types

i. RIPv1 (version 1)
   — Uses classful routing.
   — The periodic routing updates do not carry subnet information,
   — Lacking support for variable length subnet masks (VLSM).
   — There is also no support for router authentication, making RIP vulnerable to various attacks.
   — Broadcast is used for database update

ii. RIPv2 (version 2)
   — Includes the ability to carry subnet information, thus supporting Classless Inter-Domain Routing (CIDR).
   — In an effort to avoid unnecessary load on hosts that do not participate in routing,
   — RIPv2 multicasts the entire routing table to all adjacent routers at the address 224.0.0.9, as opposed to RIPv1 which uses broadcast.
   — Support Authentication

iii. RIPng (next generation)
   — Support of IPv6 networking.
   — RIPng sends updates on UDP port 521 using the multicast group FF02::9.

2. OSPF (Open Shortest Path First)

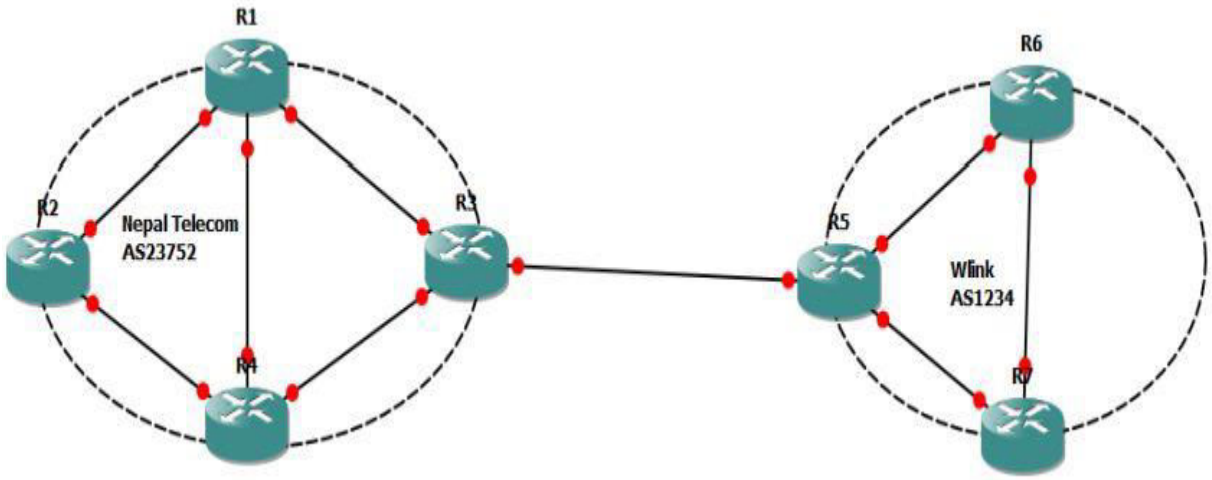
→ Link State Routing Algorithm
→ Cost/Metric = Link Bandwidth
→ Shortest Path Algorithm to calculate best path from source to destination.
→ Open Shortest Path First (OSPF) is an adaptive routing protocol for Internet Protocol (IP) networks.

Downloaded from CSIT Tutor

→ It uses a link state routing algorithm and falls into the group of interior routing protocols, operating within a single autonomous system (AS).
→ OSPF is perhaps the most widely-used interior gateway protocol (IGP) in large enterprise networks
→ It included the ability to carry subnet information, thus supporting Classless Inter-Domain Routing (CIDR).
→ Supports Authentication

3. BGP (Border Gateway Protocol)

→ Exterior Gateway protocol
→ Called Path vector Routing Algorithm.
→ Neighboring BGP routers i.e. BGP peers exchange detailed path information.
→ Used for communicating between two AS



→ Revolves around three activities
  - Receiving and filtering route advertisement from directly attached neighbors.
  - Route Selection
  - Sending route advertisements to neighbors.

ARP (Address resolution Protocol)

→ Address Resolution Protocol (ARP) is a telecommunications protocol used for resolution of network layer addresses into link layer addresses

RARP (Reverse ARP)

→ RARP (Reverse Address Resolution Protocol) is a protocol by which a physical machine in a local area network can request to learn its IP address from a gateway server's Address Resolution Protocol (ARP) table or cache.

Downloaded from CSIT Tutor

→ A network administrator creates a table in a local area network's gateway router that maps the physical machine (or Media Access Control - MAC address) addresses to corresponding Internet Protocol addresses.
→ When a new machine is set up, its RARP client program requests from the RARP server on the router to be sent its IP address. Assuming that an entry has been set up in the router table, the RARP server will return the IP address to the machine which can store it for future use.

Introduction to Multicast Routing

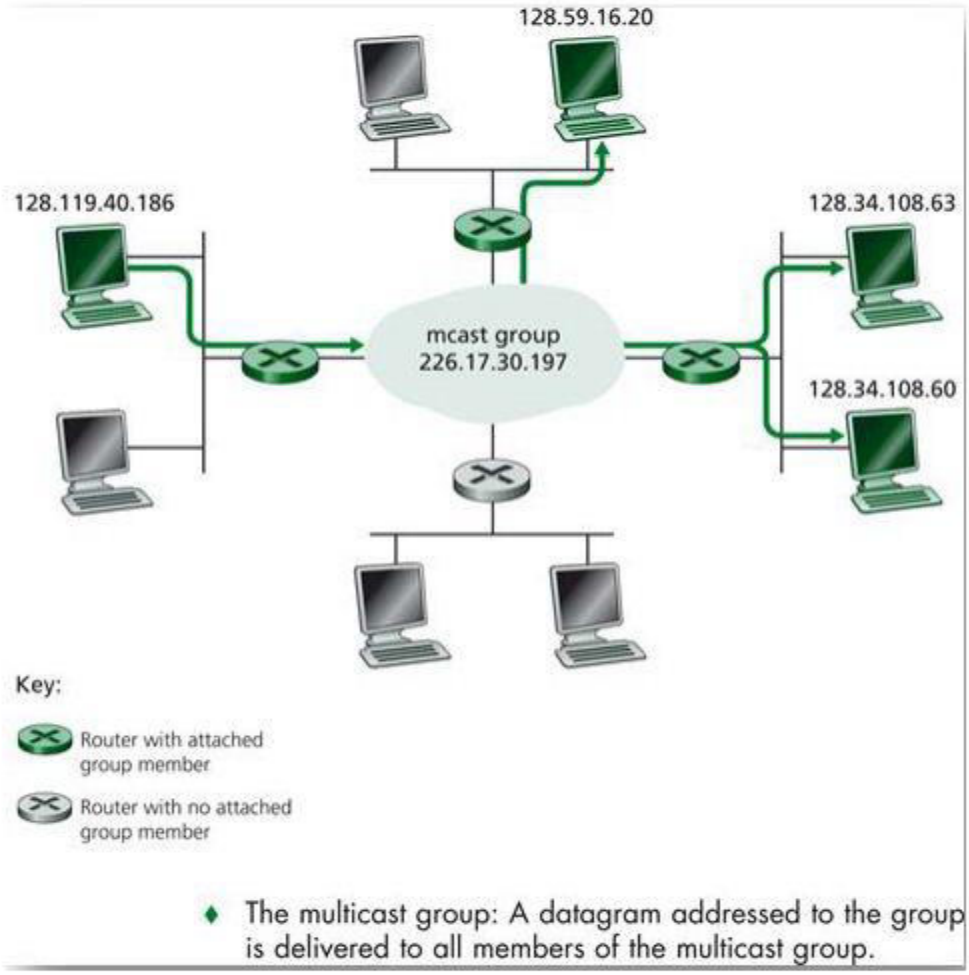
→ In computer networking, multicast is the delivery of a message or information to a group of destination computers simultaneously in a single transmission from the source.
→ Copies are automatically created in other network elements, such as routers, but only when the topology of the network requires it.
→ Multicast is most commonly implemented in IP multicast, which is often employed in Internet Protocol (IP) applications of streaming media and Internet television.
→ In IP multicast the implementation of the multicast concept occurs at the IP routing level, where routers create optimal distribution paths for datagrams sent to a multicast destination address.

Internet Group Management Protocol (IGMP)
→ IGMP runs between hosts and the nearest multicast routers.
→ A local host can use it to inform the multicast router that which multicast group it wants to be join, while the multicast routers can use it to poll the LAN periodically, thus determine if known group members are still active.

Applications of Multicast
→ Video/audio conference
→ IP TV, Video on Demand
→ Advertisement, Stock, Distance learning
→ Distributed interactive gaming or simulations
→ Voice-over-IP
→ Synchronizing of distributed database, websites

How multicast?
→ Using Class D in IP v 4 (224-239) or addresses that begin with 1111 1111 (FF) in IP v 6
→ e.g. 224.0.0.1, FF5B:2D9D:DC28:0000:0000:FC57:D4C8:1FFF
→ Rather than sending a separate copy of the data for each recipient, the source sends the data only once using the multicast group, and routers along the way to the destinations make copies as needed.

Downloaded from CSIT Tutor



IPv6

→ This huge growth in Internet use has not only led to increased demand for better, faster technology, but has also increased the demand for addresses from which to send and receive information.
→ 128 bits adresses
→ 2^{128} IP addresses developed
→ Every grain of sand on the planet can be IP-addressable

Limitations of IPv4
→ Address Space
→ Various unnecessary and Variable header fields
→ Fragmentation in Router
→ Addressing Model
→ NAT
→ Broadcast Versus Multicast

Downloaded from CSIT Tutor

→ Quality of Service

Most important changes introduced in IPv6

→ Expanded addressing capabilities
  - — Size increases from 32 bits to 128 bits. This ensures that the IP address wouldn't run out of IP addresses.
  - — In addition to unicast and multicast addresses, it introduced anycast address, which allows a datagram to be delivered to any one of a group of hosts.

→ A streamlined 40 bytes header
  - — Allows for faster processing of the IP datagram

→ Flow labeling and priority
  - — Has an elusive definition of flow.(according to quality of service or real time service e.g. audio and video transfer)

128-bit IPv6 Address

3FFE:085B:1F1F:0000:0000:0000:00A9:1234

8 groups of 16-bit hexadecimal numbers separated by ":"

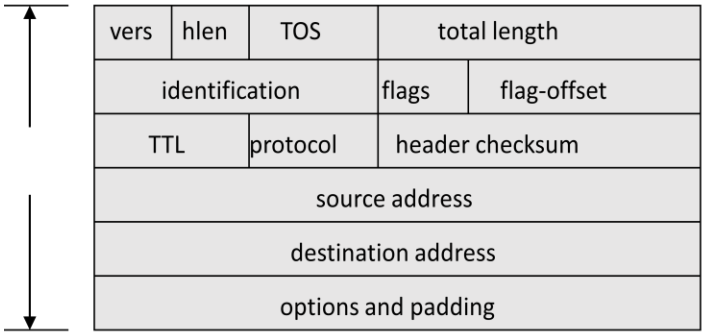
Leading zeros can be removed

3FFE:85B:1F1F::A9:1234

:: = all zeros in one or more group of 16-bit hexadecimal numbers

Header comparison

Downloaded from CSIT Tutor

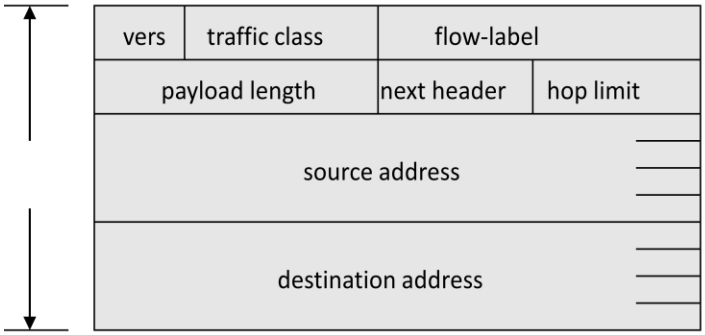


IPv4

Removed (6)

- ID, flags, flag offset
- TOS, hlen
- header checksum

Changed (3)



IPv6

Added (2)

- traffic class
- flow label

Expanded

- address 32 to 128 bits

No longer present in IPv6

→ Fragmentation/Reassembly
  — Result in fast IP forwarding

→ Header checksum:
  — Result in fast processing.

→ Option field:
  — Replaced by extension header. Result in a fixed length, 40-byte IP header.

Transition from IPv4 to IPv6

→ *Flag day* is not feasible

→ Dual stack operation – v6 nodes run in both v4 and v6 modes and use version field to decide which stack to use
  i. Nodes can be assigned a *v4 compatible v6 address*
  — Allows a host which supports v6 to talk v6 even if local routers only speak v4
  — Signals the need for tunneling
  — Add 96 0's (zero-extending) to a 32-bit v4 address – e.g. ::10.0.0.1

Downloaded from CSIT Tutor

ii. Nodes can be assigned a *v4 mapped v6 address*

— Allows a host which supports both v6 and v4 to communicate with a v4 hosts

— Add 2 bytes of 1's to v4 address then zero-extend the rest – e.g. ::ffff:10.0.0.1

→ Tunneling is used to deal with networks where v4 router(s) sit between two v6 routers

— Simply encapsulate v6 packets and all of their information in v4 packets until you hit the next v6 router.

Downloaded from CSIT Tutor

2.1 Link Layer and Local Area Networks

Introduction

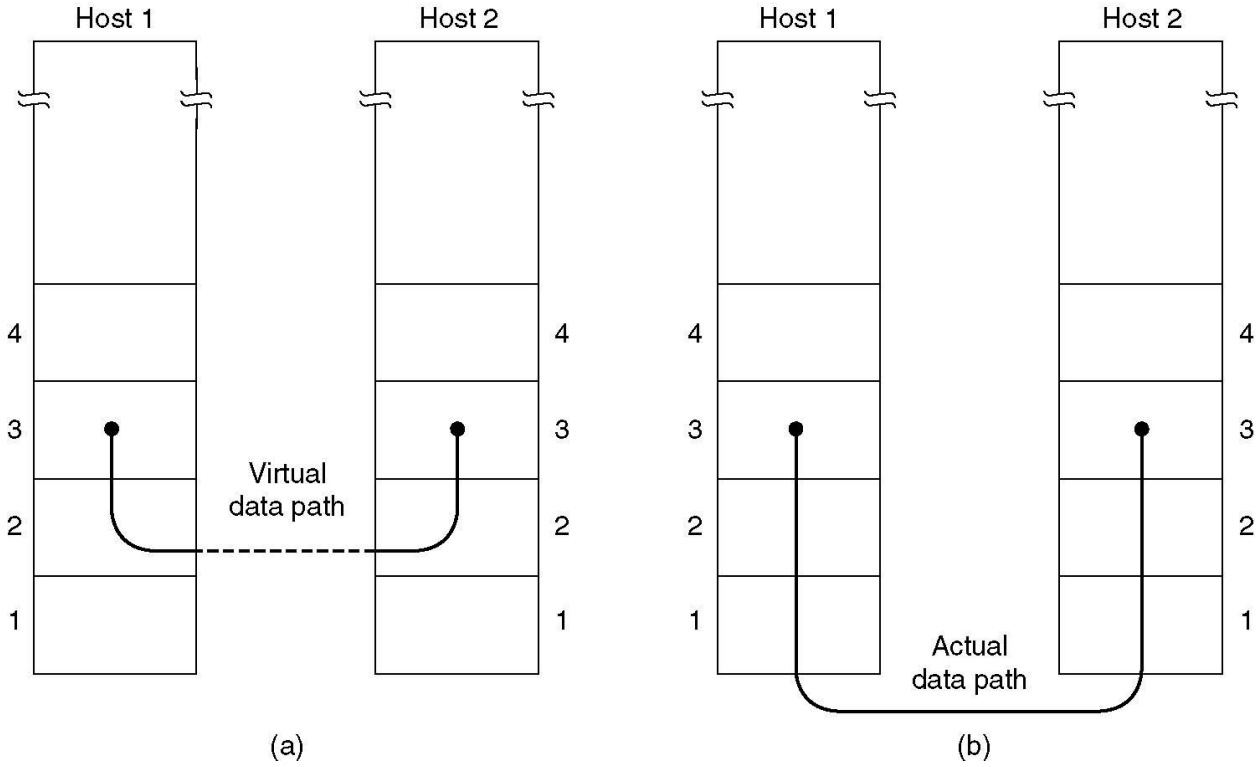
- **Data Link Layer : the services provided by the link layer**

The data link layer has three specific functions:

  - Provide a well-defined interface to the network layer.
  - Deal with transmission errors.
  - Regulate the flow of data (so that slow receivers are not overloaded).

- The Data Link Layer sits between the Network Layer and the Physical Layer.
- The DLL provides an interface for the Network Layer to send information from one machine to another.
- To the Network Layer, it looks as though the path to the new machine happens at the DLL level, when it is really happening at the physical level.

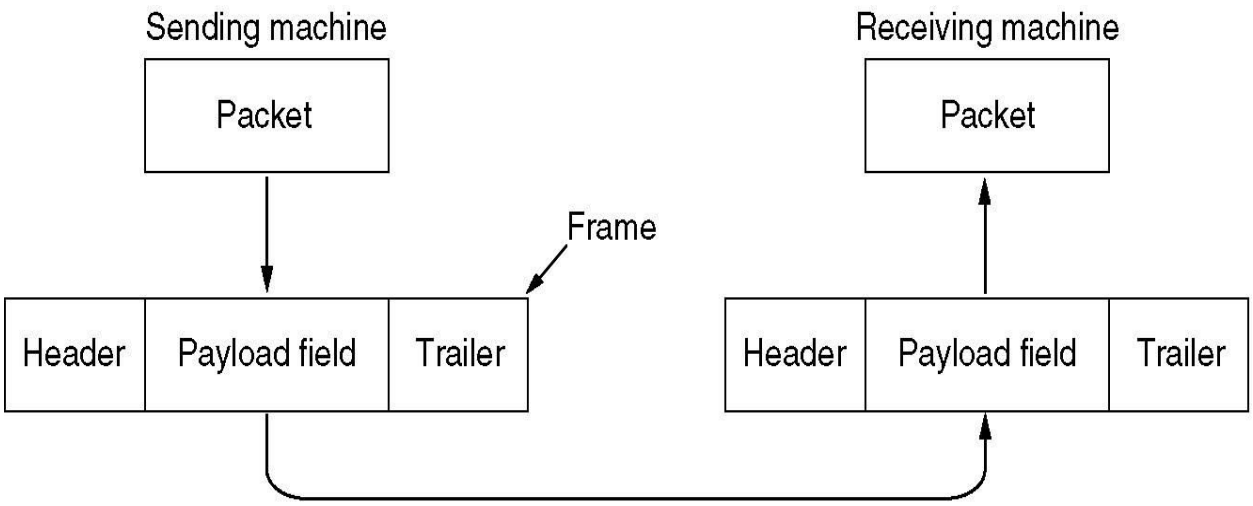
Data Flow



Link Layer Services

- Framing
  - encapsulate datagram into frame, adding header, trailer
  - The DLL is responsible for taking the **packets** of information that it receives from the Network Layer and putting them into **frames** for transmission.
  - Each frame holds the payload plus a header and a trailer (overhead).
  - It is the frames that are transmitted over the physical layer.

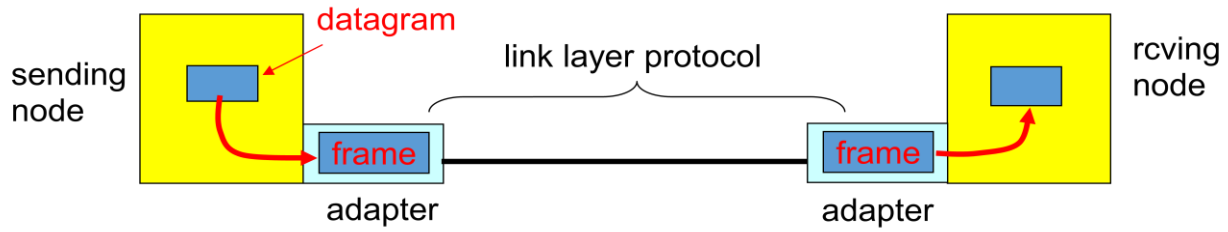
Downloaded from CSIT Tutor



- Link access
  - MAC protocol specifies the rules by which a frame is transmitted onto the link
  - channel access if shared medium
  - "MAC" addresses used in frame headers to identify source, destination
    - Different from IP address!
- Reliable delivery between adjacent nodes
  - Guarantees to move each network-layer datagram across the link without error.
  - seldom used on low bit error link (fiber, some twisted pair)
  - wireless links: high error rates
- *Flow Control:*
  - pacing between adjacent sending and receiving nodes
  - technique for speed-matching of transmitter and receiver
- *Error Detection*:
  - Errors caused by signal attenuation, noise.
  - receiver detects presence of errors:
    - signals sender for retransmission or drops frame
- Error Correction:
  - receiver identifies *and corrects* bit error(s) without resorting to retransmission
- *Half-duplex and full-duplex*
  - With full-duplex, the nodes at both ends of a link may transmit packets at the same time.
  - With half-duplex, a node cannot both transmit and receive at the same time.

Adaptors Communicating

Downloaded from CSIT Tutor



- link layer implemented in "adaptor" (aka NIC)
  - Ethernet card, 802.11 card
- sending side:
  - encapsulates datagram in a frame
  - adds error checking bits, rdt, flow control, etc.
- receiving side
  - looks for errors, rdt, flow control, etc
  - extracts datagram, passes to receiving node
- adapter is semi-autonomous
- link & physical layers

Error detection and Error correction techniques: parity checks, checksum, CRC

Error Detection with the CRC

CRC (cyclical redundancy checking) is a rather sophisticated *error checking* routine used to assure that two communicating computers are really getting the correct data. There are a whole bunch of these error correction and checking methods out there, but they all have pretty much the same objective: making sure that no data is lost or misinterpreted when two computers are exchanging information over the phone lines (through a modem, of course).

The Cyclic Redundancy Check (CRC)

Consider a message having many data bits which are to be transmitted reliably by appending several check bits as shown below.

message bits	parity bits

The exact number of extra bits and their makeup depends on a generating polynomial. For example one such polynomial is:

$$x^{16} + x^{12} + x^5 + 1$$

A Standard Generating Polynomial - CRC(CCITT)

The number of CRC bits corresponds to the order of the generating polynomial. The above polynomial of order 16 generates a 16-bit CRC. Typically, the CRC bits are used for error *detection* only.

CRC Computation

Downloaded from CSIT Tutor

Consider a message represented by some polynomial M(x), and a generating polynomial G(x).
In this example, let M(x) represent the binary message 110010, and let G(x) = $x^3 + x^2 + 1$; (binary **1101**).
The polynomial G(x) will be used to generate a (3-bit) CRC called R(x) which will be appended to M(x). Note that G(x) is prime.

110010	CRC

The polynomial G(x) defines the CRC bits.

Step 1 - Multiply the message M(x) by x^3, where 3 is the number of bits in the CRC.
Add three zeros to the binary M(x).
Step 2 - Divide the product x^3 [M(x)] by the generating polynomial G(x).
We wish to find "the remainder, modulo M(x)"
Compute the following:

```
        100100 (ignore this quotient)
      ------------
1101) 110010000
      1101
      ----
        1100
        1101
        ----
         100   = remainder = R(x)
```

Observe that if R(x) were in place of the appended zeros, the remainder would become 000.
Step 3 - Disregard the quotient and add the remainder R(x)to the product x^3 [M(x)] to yield the code message polynomial T(x), which is represented as:
$T(x) = x^3$ [M(x)] + R(x)
Put the remainder R(x)=100 in place of the three zeros added in Step 1.

110010	100

The message may now be transmitted

CRC Error Checking - No Errors
Upon reception, the entire received T(x) = "message + crc" can be checked simply by dividing T(x)/G(x) using the same generating polynomial. If the remainder after division equals zero, then no error was found.

```
        100100 (ignore this quotient)
      ------------
1101) 110010100
      1101
      ----
        1101
```


Downloaded from CSIT Tutor

```
        1101
        ----
         000   = remainder (no error)
```

CRC Error Checking - Single Bit Error

A single bit error in bit position K in a message T(x) can be represented by adding the term E(x) = x^K, (binary 1 followed by K-zeros).

```
          sent: 110010100  =  T(x)
         error: 000001000  =  E(x) = x^3
                ---------
      received: 110011100  =  T(x) + E(x)    (error in bit 3)
```

The above error would be detected when the CRC division is performed:

```
          100101 (ignore this quotient)
       ------------
  1101) 110011100 = T(x) + E(x)
        1101
        ----
         1111
         1101
         ----
          1000
          1101
          ----
           101   = remainder (error!)
```

Note that division by G(x) revealed the error. On the other hand, since T(x)/G(x) = 0 by definition, the remainder is a function only of the error. An error in this same bit would give the same non-zero remainder *regardless of the message bits*.

$$\frac{T(x)+E(x)}{G(x)} = \frac{T(x)}{G(x)} + \frac{E(x)}{G(x)} = \frac{E(x)}{G(x)}$$

The remainder is a function only of the error bits E(x).

```
          1 (ignore this quotient)
       ------------
  1101) 000001000   = E(x) alone
        1101
        ----
         101   = remainder (error!)
```

Since E(x) = x^K has no factors other than x, a single bit error will never produce a term exactly divisible by G(x). **All single bit errors will be detected.**

Downloaded from CSIT Tutor

Multiple Access Protocols (CHANNEL-PARTITIONING, RANDOM ACCESS, TAKING-TURNS)

- single shared broadcast channel
- two or more simultaneous transmissions by nodes:
  Interference
  - collision: if node receives two or more signals at the same time

Multiple access protocol

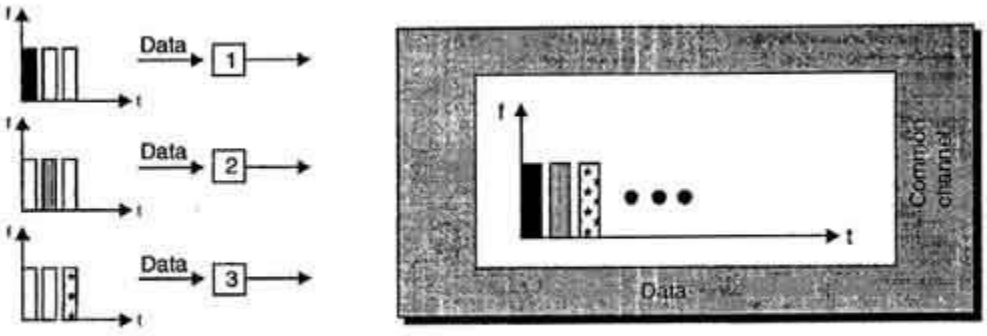
- distributed algorithm that determines how nodes share channel, i.e., determine when node can transmit.
- communication about channel sharing must use channel itself!
  - no out-of-band channel for coordination

CHANNEL PARTITIONING PROTOCOL

I. **TDMA** (Time Division Multiple Access)

- In TDMA, the bandwidth of channel is dividend amongst various stations on the basis of time.
- Each station is allocated a time slot during which it can send its data *i.e.* each station can transmit its data in its allocated time slot only.
- Each station must know the beginning of its slot and the location of its slot.
- TDMA requires synchronization between different stations.
- Synchronization is achieved by using some synchronization bits (preamble bits) at the beginning of each slot.
- TDMA is different from TDM, although they are conceptually same.
- TDM is a physical layer technique that combines the data from slower channels and transmits then by using a faster channel. This process uses physical multiplexer.
- TDMA, on other hand, is an access method in the data link layer. The data link layer in each station tells its physical layer to use the allocated time slot. There is no physical multiplexer at the physical layer.

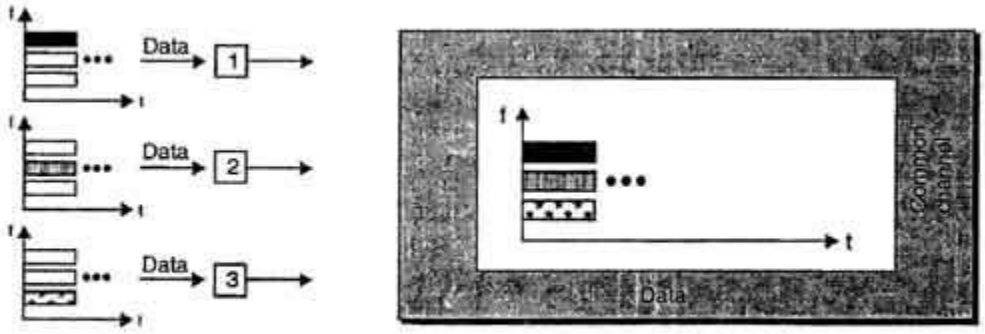
Downloaded from CSIT Tutor



TDMA

II. FDMA (Frequency Division Multiple Access)

- In FDMA, the available bandwidth is divided into various frequency bands.
- Each station is allocated a band to send its data. This band is reserved for that station for all the time.
- The frequency bands of different stations are separated by small bands of unused frequency. These unused frequency bands are called guard bands that prevent station interferences.
- FDMA is different from frequency division multiplexing (FDM).
- FDM is a physical layer technique whereas FDMA is an access method in the data link layer.
- FDM combines loads from different low bandwidth channels and transmit them using a high bandwidth channel. The channels that are combined are low-pass. The multiplexer modulates the signal, combines them and creates a band pass signal. The bandwidth of each channel is shifted by the multiplexer.
- In FDMA, data link layer in each station tells its physical layer to make a band pass signal from the data passed to it. The signal must be created in the allocated band. There is no physical multiplexer at the physical layer.



FDMA

III. **CDMA** (Code Division Multiple Access)

Downloaded from CSIT Tutor

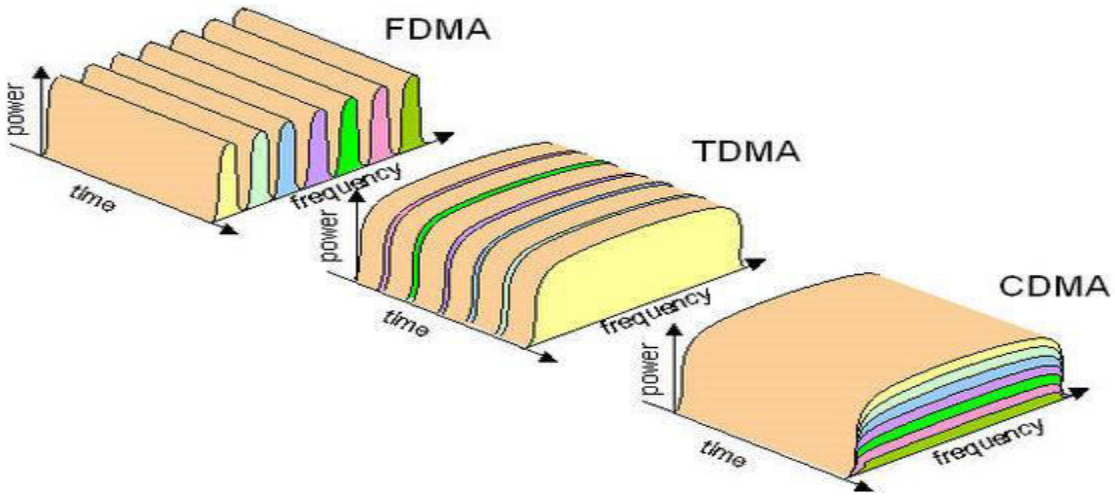
CDMA (Code Division Multiple Access) also called *spread-spectrum* and *code division multiplexing,* one of the competing transmission technologies for digital MOBILE PHONES. The transmitter mixes the packets constituting a message into the digital signal stream in an order determined by a PSEUDO-RANDOM NUMBER sequence that is also known to the intended receiver, which uses it to extract those parts of the signal intended for itself. Hence, each different random sequence corresponds to a separate communication channel. CDMA is most used in the USA.

• Unlike TDMA, in CDMA all stations can transmit data simultaneously, there is no timesharing.

• CDMA allows each station to transmit over the entire frequency spectrum all the time.

• Multiple simultaneous transmissions are separated using coding theory.

• In CDMA, each user is given a unique code sequence.



RANDOM ACCESS PROTOCOL

I. Aloha

ALOHA: ALOHA is a system for coordinating and arbitrating access to a shared communication Networks channel. It was developed in the 1970s by Norman Abramson and his colleagues at the University of Hawaii. The original system used for ground based radio broadcasting, but the system has been implemented in satellite communication systems.

A shared communication system like ALOHA requires a method of handling collisions that occur when two or more systems attempt to transmit on the channel at the same time. In the ALOHA system, a node transmits whenever data is available to send. If another node transmits at the same time, a collision occurs, and the frames that were transmitted are lost. However, a

Downloaded from CSIT Tutor

node can listen to broadcasts on the medium, even its own, and determine whether the frames were transmitted.

Aloha means "Hello". Aloha is a multiple access protocol at the data link layer and proposes how multiple terminals access the medium without interference or collision. In 1972, Roberts developed a protocol that would increase the capacity of aloha two fold. The Slotted Aloha protocol involves dividing the time interval into discrete slots and each slot interval corresponds to the time period of one frame. This method requires synchronization between the sending nodes to prevent collisions.

There are two different versions/types of ALOHA:

a) **Pure Aloha**

• **In** pure ALOHA, the stations transmit frames whenever they have data to send.

• When two or more stations transmit simultaneously, there is collision and the frames are destroyed.

• In pure ALOHA, whenever any station transmits a frame, it expects the acknowledgement from the receiver.

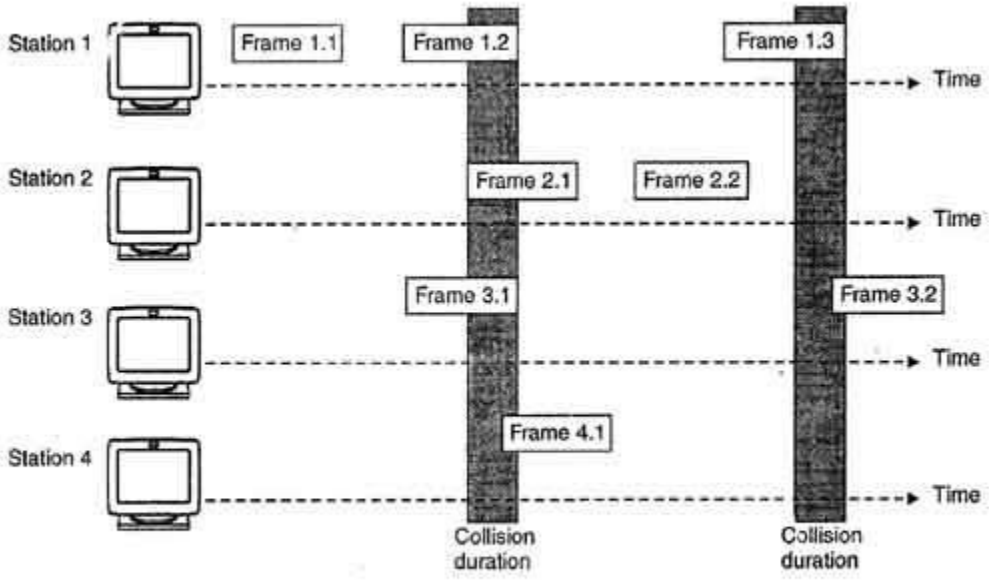
• If acknowledgement is not received within specified time, the station assumes that the frame (or acknowledgement) has been destroyed.

• If the frame is destroyed because of collision the station waits for a random amount of time and sends it again. This waiting time must be random otherwise same frames will collide again and again.

• Therefore, pure ALOHA dictates that when time-out period passes, each station must wait for a random amount of time before resending its frame. This randomness will help avoid more collisions.

• Figure shows an example of frame collisions in pure ALOHA.

Downloaded from CSIT Tutor



Frames in Pure ALOHA

• In figure, there are four stations that contended with one another for access to shared channel. All these stations are transmitting frames. Some of these frames collide because multiple frames are in contention for the shared channel. Only two frames, frame 1.1 and frame 2.2 survive. All other frames are destroyed.

• Whenever two frames try to occupy the channel at the same time, there will be a collision and both will be damaged. If first bit of a new frame overlaps with just the last bit of a frame almost finished, both frames will be totally destroyed and both will have to be retransmitted.

b) Slotted Aloha

• Slotted ALOHA was invented to improve the efficiency of pure ALOHA as chances of collision in pure ALOHA are very high.

• In slotted ALOHA, the time of the shared channel is divided into discrete intervals called slots.

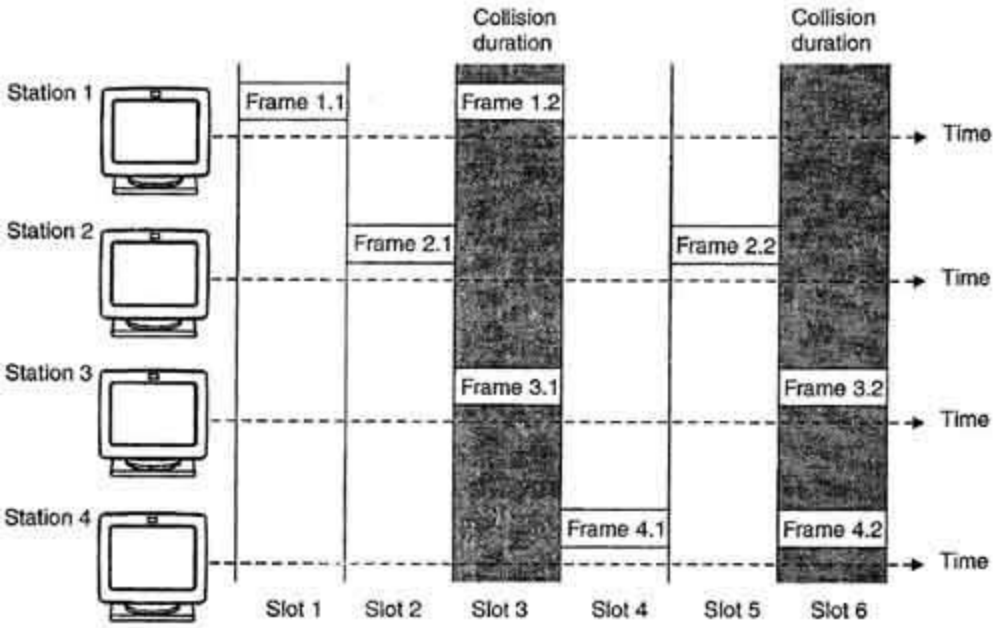
• The stations can send a frame only at the beginning of the slot and only one frame is sent in each slot.

• In slotted ALOHA, if any station is not able to place the frame onto the channel at the beginning of the slot *i.e.* it misses the time slot then the station has to wait until the beginning of the next time slot.

• In slotted ALOHA, there is still a possibility of collision if two stations try to send at the beginning of the same time slot as shown in figure.

• Slotted ALOHA still has an edge over pure ALOHA as chances of collision are reduced to one-half.

Downloaded from CSIT Tutor



Frames in Slotted ALOHA

Suppose there is only one channel and two computers C1 and C2 are willing to send data through it. If C1 is transmitting data, it sends signals to all the other computers notifying that C1 is about to send data. After the time slot has completed, only then C2 can transmit data through that channel.

II. CSMA

Carrier Sense Multiple Access (CSMA): CSMA is a network access method used on shared network topologies such as Ethernet to control access to the network. Devices attached to the network cable listen (carrier sense) before transmitting. If the channel is in use, devices wait before transmitting. MA (Multiple Access) indicates that many devices can connect to and share the same network. All devices have equal access to use the network when it is clear.

CSMA protocol was developed to overcome the problem found in ALOHA i.e. to minimize the chances of collision, so as to improve the performance. CSMA protocol is based on the principle of 'carrier sense'. The station senses the carrier or channel before transmitting a frame. It means the station checks the state of channel, whether it is idle or busy.

Even though devices attempt to sense whether the network is in use, there is a good chance that two stations will attempt to access it at the same time. On large networks, the transmission time between one end of the cable and another is enough that one station may access the cable even though another has already just accessed it.

Downloaded from CSIT Tutor

The chances of collision still exist because of propagation delay. The frame transmitted by one station takes some time to reach other stations. In the meantime, other stations may sense the channel to be idle and transmit their frames. This results in the collision.

Consider computers C1, C2, and C3 are willing to send data through a channel. At first, C1 transmits data. In the due course of transmission, C2 and C3 check the status of the channel at the same time. Both find the channel to be busy, so they wait for time T. After time T, both C2 and C3 check the channel and find it free, so they start to initiate the process of data transmission which can lead to collision.

CSMA modes:

1-persistent

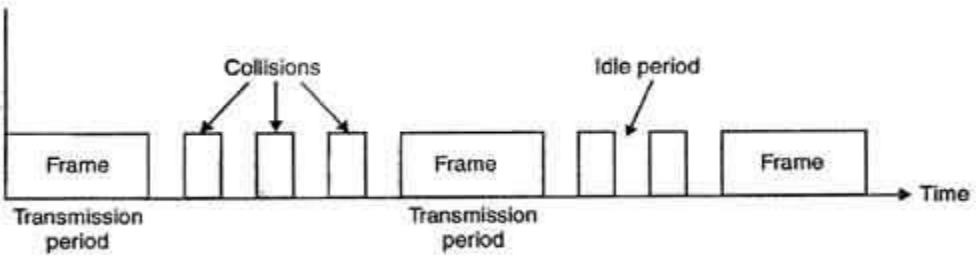
Non-persistent

P-persistent

a) CSMA/CD

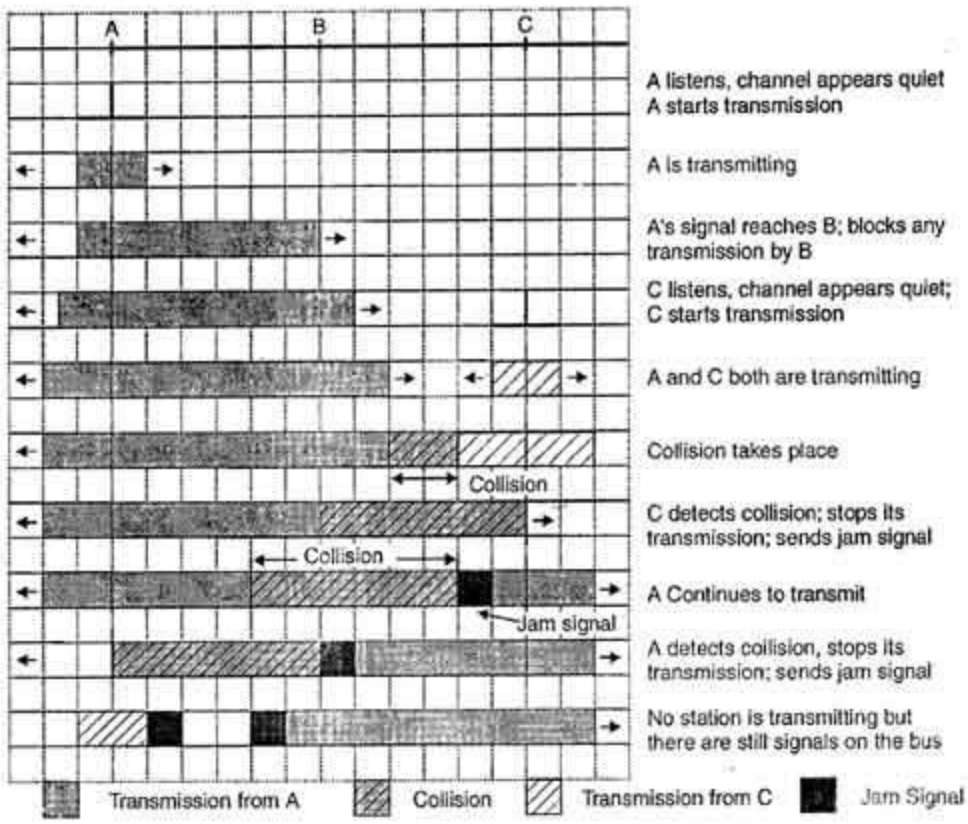
- *CSMA/CD* is a protocol in which the station senses the carrier or channel before transmitting frame just as in persistent and non-persistent CSMA. If the channel is busy, the station waits.
- Additional feature in *CSMA/CD* is that the stations can detect the collisions. The stations abort their transmission as soon as they detect a collision. In CSMA, this feature is not present. The stations continued their transmission even though they find that the collision has occurred. This leads to the wastage of channel time.
- However, this problem is handled in *CSMA/CD*. In *CSMA/CD,* the station that places its data onto the channel after sensing the channel continues to sense the channel even after the data transmission. If collision is detected, the station aborts its transmission and waits for predetermined amount of time & then sends its data again.
- As soon as a collision is detected, the transmitting station releases a jam signal.
- Jam signal will alert the other stations. The stations are not supposed to transmit immediately after the collision has occurred. Otherwise, there is a possibility that the same frames would collide again.
- After some back-off delay time, the stations will retry the transmission. If the collision occurs again then the back-off delay time is increased progressively.
- Therefore, the CSMA/CD method consists of alternating transmission period and collisions with idle periods when none of the stations is transmitting.

Downloaded from CSIT Tutor



CSMA/CD with three states : collisions, transmission, or idle

The entire scheme of CSMA/CD is depicted in the figure:



CSMA/CD

i. IEEE 802.3 Ethernet

The frame format specified by IEEE 802.3 standard contains following fields:

Preamble	Destination Address	Source Address	Type	Data	Frame Check Status (FCS)
Bytes 8	6	6	2	46 to 1500 bytes	2 or 4

Preamble: It is seven bytes (56 bits) that provides bit synchronization. It consists of alternating 0s and 1s. The purpose is to provide alert and timing pulse.

Downloaded from CSIT Tutor

Destination Address (DA): It is six byte field that contains physical address of packet's destination.

Source Address (SA): It is also a six byte field and contains the physical address of source or last device to forward the packet (most recent router to receiver).

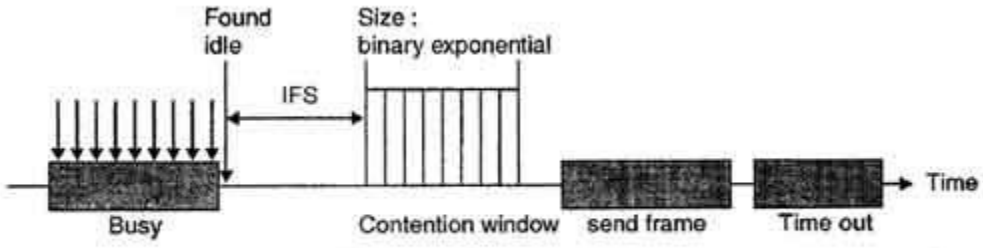
Length: This two byte field specifies the length or number of bytes in data field.

Data: It can be of 46 to 1500 bytes, depending upon the type of frame and the length of the information field.

Frame Check Sequence (FCS): This is for byte field, contains CRC for error detection.

b) CSMA/CA

- CSMA/CA protocol is used in wireless networks because they cannot detect the collision so the only solution is collision avoidance.
- CSMA/CA avoids the collisions using three basic techniques.
  a. Interframe space
  b. Contention window
  c. Acknowledgements



CSMA/CA

a. Interframe Space (IFS)

• Whenever the channel is found idle, the station does not transmit immediately. It waits for a period of time called interframe space (IFS).

• When channel is sensed to be idle, it may be possible that same distant station may have already started transmitting and the signal of that distant station has not yet reached other stations.

• Therefore the purpose of IFS time is to allow this transmitted signal to reach other stations.

• If after this IFS time, the channel is still idle, the station can send, but it still needs to wait a time equal to contention time.

• IFS variable can also be used to define the priority of a station or a frame.

b. Contention Window

• Contention window is an amount of time divided into slots.

• A station that is ready to send chooses a random number of slots as its wait time.

Downloaded from CSIT Tutor

• The number of slots in the window changes according to the binary exponential back-off strategy. It means that it is set of one slot the first time and then doubles each time the station cannot detect an idle channel after the IFS time.
• This is very similar to the p-persistent method except that a random outcome defines the number of slots taken by the waiting station.
• In contention window the station needs to sense the channel after each time slot.
• If the station finds the channel busy, it does not restart the process. It just stops the timer & restarts it when the channel is sensed as idle.

c. Acknowledgement

• Despite all the precautions, collisions may occur and destroy the data.
• The positive acknowledgment and the time-out timer can help guarantee that receiver has received the frame.

i. IEEE 802.11 Wireless LAN (WiFi)

Wireless communication is one of the fastest growing technologies these days. Wireless LANs are commonly found in office buildings, college campuses, and in many public areas.
Types

Standard	**Frequency Range (US)**	**Data Rate**
IEEE 802.11b	2.4 – 2.485 GHz	Up to 11 Mbps
IEEE 802.11a	5.1 – 5.8 GHz	Up to 54 Mbps
IEEE 802.11g	2.4 – 2.485 GHz	Up to 54 Mbps

IEEE 802.11n is on the process of standardization, uses Multiple Input Multiple Output (MIMO) antennas.

IEEE 802.11 standard provides wireless communication with the use of infrared or radio waves.

- Two configurations:

➢ Ad-hoc: no central control, no connection to the outside world
➢ Infrastructure: uses fixed network access point to connect to the outside world.

- It doesn't implement collision detection because it can't detect collisions at the receiver end (hidden terminal problem)
- To avoid collisions, the frames contains field containing the length of the transmissions. Other stations defer transmissions.
- 802.11 lives in physical layer and data link layer in the OSI.

Downloaded from CSIT Tutor

- IEEE 802.11b (Wi-Fi) is a wireless LAN technology that is growing rapidly in popularity. It is convenient, inexpensive and easy to use.
  Uses: airports, hotels, bookstores, parks etc.
  Estimates: 70% of WLANs are insecure.
- 802.11b has a maximum raw data rate of 11 Mbit/s and uses the same media access method defined in the original standard. 802.11b products appeared on the market in early 2000, since 802.11b is a direct extension of the modulation technique defined in the original standard. The dramatic increase in throughput of 802.11b (compared to the original standard) along with simultaneous substantial price reductions led to the rapid acceptance of 802.11b as the definitive wireless LAN technology.
- 802.11b devices experience interference from other products operating in the 2.4 GHz band. Devices operating in the 2.4 GHz range include microwave ovens, Bluetooth devices, baby monitors, cordless telephones and some amateur radio equipment.

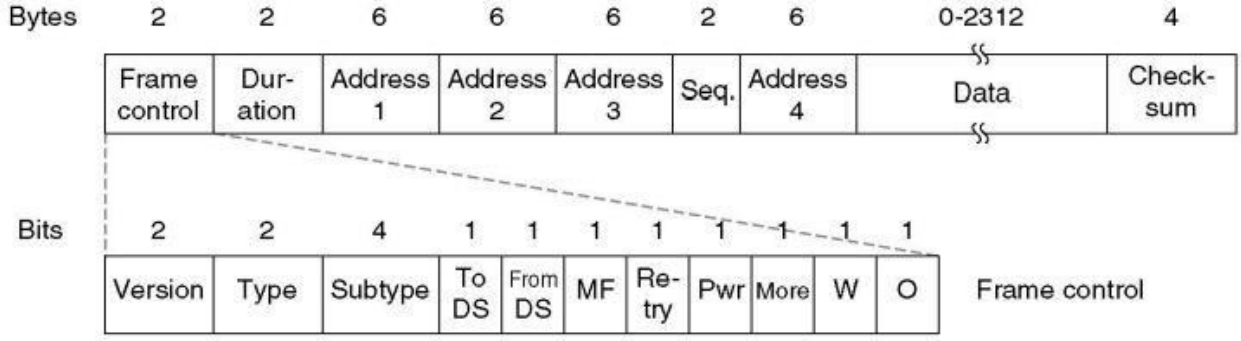


Fig: the 802.11 frame structure

Frame Control: Contains following

- Version: Protocol version Type: data, control or mgmt. Subtype : RTS or CTS
- To/From DS: Going to or Coming from intercell distribution (e.g. Ethernet)
- MF: More fragments to follow
- Retry: Retransmission of earlier frame
- Pwr: used by base station to sleep or wake receiver
- More: sender has more frames for receiver
- W: WEP Encryption
- O : sequence of frames must be processed in order

Duration: time to occupy channel, used by other stations to manage NAV

Addresses: Two are source and destination. Add, of sender and receiver, other two are that of base stations for intercell traffic.

TAKING-TURNS PROTOCOL

I. Polling Protocol

Downloaded from CSIT Tutor

The polling protocol requires one of the nodes to be designated as a master node which polls each of the nodes in a round-robin fashion. In particular, the master node first sends a message to node 1, saying that it (node 1) can transmit up to some maximum number of frames. After node 1 transmits some frames, the master node tells node 2 that it (node 2) can transmit up to the maximum number of frames. (The master node can determine when a node has finished sending its frames by observing the lack of a signal on the channel.) The procedure continues in this manner, with the master node polling each of the nodes in a cyclic manner.

a) Bluetooth

Bluetooth is a wireless technology standard for exchanging data over short distances (using short-wavelength UHF radio waves in the ISM band from 2.4 to 2.485 GHz) from fixed and mobile devices, and building personal area networks (PANs). Invented by telecom vendor Ericsson in 1994, it was originally conceived as a wireless alternative to RS-232 data cables. It can connect several devices, overcoming problems of synchronization.

Bluetooth is managed by the Bluetooth Special Interest Group (SIG), which has more than 20,000 member companies in the areas of telecommunication, computing, networking, and consumer electronics. Bluetooth was standardized as IEEE 802.15.1, but the standard is no longer maintained. The SIG oversees the development of the specification, manages the qualification program, and protects the trademarks. To be marketed as a Bluetooth device, it must be qualified to standards defined by the SIG. A network of patents is required to implement the technology, which is licensed only for that qualifying device.

II. Token-passing Protocol

This protocol, also called Token-Passing Protocol, relies on a control signal called the token. A token is a 24-bit packet that circulates throughout the network from NIC to NIC in an orderly fashion. If a workstation wants to transmit a message, first it must seize the token. At that point, the workstation has complete control over the communications channel. The existence of only one token eliminates the possibility of signal collisions. This means that only one station can speak at a time.

a) IEEE 802.5 Token Ring

Ring technology is a collection of point-to-point links that happen to a form of circle, not a broadcast medium, which supports to run twisted pair, coax and fiber optic cables. Each bit arriving at an interface of the ring is copied into a 1-bit buffer and then copied out onto the ring again. While in the buffer, the bit can be inspected and possibly modified before being written out. This copying step introduces 1-bit delay at each interface.

Downloaded from CSIT Tutor

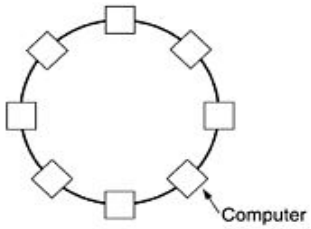


Fig: Token Ring

In token ring special bit pattern, called the token, circulates around the ring whenever all stations are idle. When a station wants to transmit a frame, it is required to seize the token and remove it from the ring before transmitting. This action is done by inverting a single bit in the 3 byte token, which instantly changes it into the first 3 bytes of normal data. Because there is only one token, only one station can transmit at a given instant, thus solving the channel access problem the same way token bus solves it.

A station may hold the token for the token holding time, which is 10 ms unless an installation sets a different value. After all frames transmitted or the transmission of another frame would exceed the token holding time, the station regenerates the token.

Token ring frame format:

	SD	AC	FC	DA	SA	Data	Checksum	ED	FS
Bytes	1	1	1	2 or 6	2 or 6	>=0	4	1	1

Data/ Command

Start Delimiter	Access Control	End Delimiter
1	1	1

Token

Token Frame Fields

• Start delimiter—Alerts each station of the arrival of a token (or data/command frame). This field includes signals that distinguish the byte from the rest of the frame by violating the encoding scheme used elsewhere in the frame.

• Access-control byte—Contains the Priority field (the most significant 3 bits) and the Reservation field (the least significant 3 bits), as well as a token bit (used to differentiate a token from a data/command frame) and a monitor bit (used by the active monitor to determine whether a frame is circling the ring endlessly).

Downloaded from CSIT Tutor

• End delimiter—Signals the end of the token or data/command frame. This field also contains bits to indicate a damaged frame and identify the frame that is the last in a logical sequence.

Data/Command Frame Fields

• Start delimiter—Alerts each station of the arrival of a token (or data/command frame). This field includes signals that distinguish the byte from the rest of the frame by violating the encoding scheme used elsewhere in the frame.

• Access-control byte—Contains the Priority field (the most significant 3 bits) and the Reservation field (the least significant 3 bits), as well as a token bit (used to differentiate a token from a data/command frame) and a monitor bit (used by the active monitor to determine whether a frame is circling the ring endlessly).

• Frame-control bytes—Indicates whether the frame contains data or control information. In control frames, this byte specifies the type of control information.

• Destination and source addresses—Consists of two 6-byte address fields that identify the destination and source station addresses.

• Data—Indicates that the length of field is limited by the ring token holding time, which defines the maximum time a station can hold the token.

• Frame-check sequence (FCS)—Is filed by the source station with a calculated value dependent on the frame contents. The destination station recalculates the value to determine whether the frame was damaged in transit. If so, the frame is discarded.

• End Delimiter—Signals the end of the token or data/command frame. The end delimiter also contains bits to indicate a damaged frame and identify the frame that is the last in a logical sequence.

• Frame Status—Is a 1-byte field terminating a command/data frame. The Frame Status field includes the address-recognized indicator and frame-copied indicator.

b) FDDI

- Fiber Distributed Data Interface
- Similar to Token ring in the sense that it share some features such as topology(ring) and media access technique(token-passing)
- High performance Fiber Optic token ring running at 100 mbps over distance 200 KM and permits up to 1000 stations
- FDDI deals with network reliable issues as mission-critical applications were implemented on high speed networks. It is frequently used as a backbone technology, and to connect high speed computer on LAN
- Based on two counter–rotating fiber rings, only one used at a time and next is for backup. So if there is any problem in one ring, next ring works automatically
- It allows 16 to 48 bits address and maximum frame size is 4500 bytes
- It prefers multimode fiber optic cable rather than single mode as multimode reduces cost for high data transmission

Downloaded from CSIT Tutor

- It prefers LEDs instead of Laser for light source not only for cheaper but also to remove accidental chances at user end connector (if user open connector and sees cable by naked eye, eye may damage on laser light)
- It operates at low error (1 bit error for 2.5 x 10^{10})
- It uses 4B/5B encoding in place of Manchester encoding in Token Ring
- It capture token before transmitting and does not wait for acknowledgement to regenerate token as ring might be very long and may occurs much delay to wait for ACK.
- In normal operation, the token and frames travel only on the primary ring in a single direction. The second ring transmits idle signals in the opposite direction
- If a cable or device becomes disabled, the primary ring raps back around onto the secondary ring
- Stations may be directly connected to FDDI dual ring or attached to FDDI concentrator. There are three types of nodes:
  - -DAS (Dual attachment station)
  - -SAS (Single attachment station)
  - -DAC (Dual attachment concentrator)
- FDDI deploys following timers:
  - -Token holding time: upper limit on how long a station can hold token
  - -Token Rotation time: how long it takes the token to traverse the ring or the interval between two successive arrivals of the token
- There are four specifications in FDDI.
  - *Media Access control*- deals with how medium is accessed, frame format, token handling, addressing, fair and equal access of the ring through the use of the timed token, guarantee bandwidth for special traffic etc.
  - *Physical layer protocol*-deals with data encoding/decoding procedures, establish clock synchronization, data recovery from incoming signal etc.
  - *Physical layer medium*- defines characteristics of transmission medium, fiber optic link type: single mode, multimode; power levels, bit error rates, optical components: connectors, switches, LEDs, Pin etc.
  - *Station Management*- defines FDDI station configuration, ring configuration, ring control features, station insertion and removal, initialization etc.

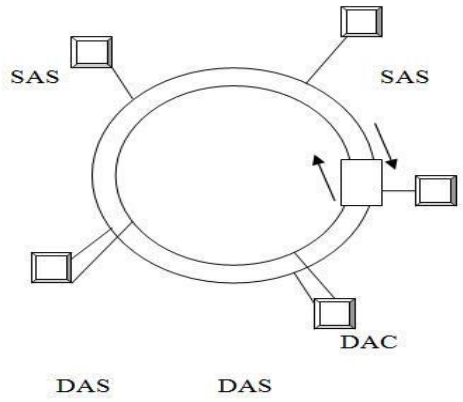


Fig: FDDI Dual Ring

Downloaded from CSIT Tutor

FDDI Frame format:

Preamble	SD	FC	DA	SA	Data	Checksum	ED	FS
8 B	1 B	1 B	2 or 6 B	2or 6B	4500 B	4 B	1 B	1B

FDDI Frame can be as long as 4500bytes.
Preamble: Unique sequence that prepares each station for an upcoming frame.
Start Delimiter: Indicates beginning of the frame.
Frame Control: Indicates size of address field and whether the frame contains synchronous or asynchronous data, among other control information
Destination Address: Contains a unicast, multicast or broadcast address. FDDI uses 6 byte address
Source Address: 6 byte address source Address.
Data: Contains either information destined for upper layers or control information
Frame Check Sequence: For Error detection.
End Delimiter: End of Frame.
Frame status: Allows the source station to determine whether an error occurred; identifies whether the frame was recognized and copied by a receiving station.

ARP

- Address Resolution Protocol
- Used to convert an IP address into a physical address (called a *DLC address*), such as an Ethernet address.

RARP

- Reverse ARP
- used by a host to discover its IP address
- to convert physical address into IP address

PPP – the Point-to-Point Protocol

- PPP was devised by IETF (Internet Engineering Task Force) to create a data link protocol for point to point lines that can solve all the problems present in SLIP (serial line internet protocol).
- PPP is most commonly used data link protocol. It is used to connect the Home PC to the server of ISP via a modem.

This protocol offers several facilities that were not present in SLIP. Some of these facilities are:

i. PPP defines the format of the frame to be exchanged between the devices.
ii. It defines link control protocol (LCP) for:-
    (a) Establishing the link between two devices.
    (b) Maintaining this established link.

Downloaded from CSIT Tutor

(c) Configuring this link.

(d) Terminating this link after the transfer.

iii. It defines how network layer data are encapsulated in data link frame.

iv. PPP provides error detection.

v. Unlike SLIP that supports only IP, PPP supports multiple protocols.

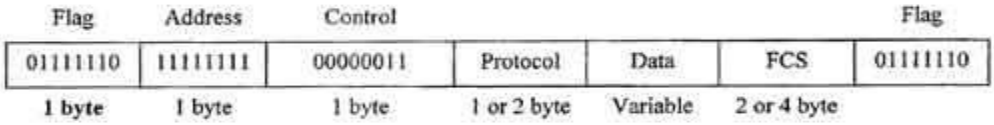
vi. PPP allows the IP address to be assigned at the connection time i.e. dynamically. Thus a temporary IP address can be assigned to each host.

vii. PPP provides multiple network layer services supporting a variety of network layer protocol. For this PPP uses a protocol called NCP (Network Control Protocol).

viii. It also defines how two devices can authenticate each other.

PPP Frame Format

The frame format of PPP resembles HDLC frame. Its various fields are:



PPP frame format

Flag field: Flag field marks the beginning and end of the PPP frame. Flag byte is 01111110. (1 byte).

Address field: This field is of 1 byte and is always 11111111. This address is the broadcast address *i.e.* all the stations accept this frame.

Control field: This field is also of 1 byte. This field uses the format of the U-frame (unnumbered) in HDLC. The value is always 00000011 to show that the frame does not contain any sequence numbers and there is no flow control or error control.

Protocol field: This field specifies the kind of packet in the data field *i.e.* what is being carried in data field.

Data field: Its length is variable. If the length is not negotiated using LCP during line set up, a default length of 1500 bytes is used. It carries user data or other information.

FCS field: The frame checks sequence. It is either of 2 bytes or 4 bytes. It contains the checksum.

ATM

Asynchronous Transfer Mode (ATM) is also called *cell relay,* a high-speed switched network technology developed by the telecommunications industry to implement the next, BROADBAND generation of ISD. ATM was designed for use in WANS such as the public telephone system and corporate data networks, though it has also been applied to create super-

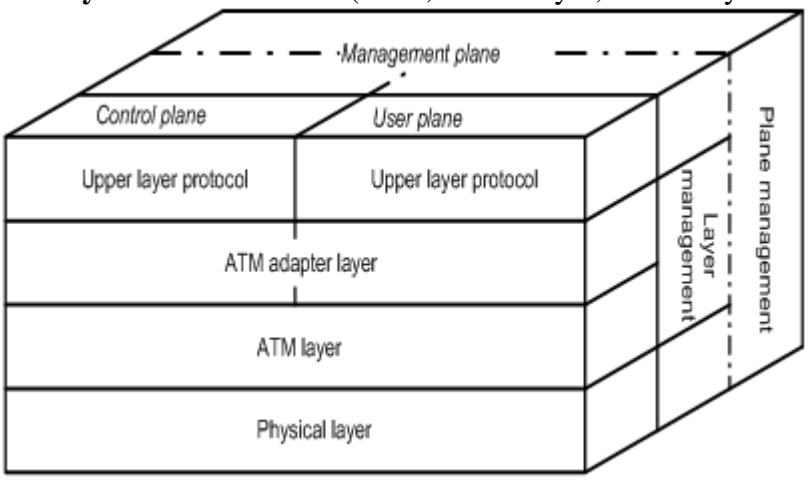
Downloaded from CSIT Tutor

fast LANS. It can carry all kinds of traffic - voice, video and data – simultaneously at speeds up to 155megabits per second.

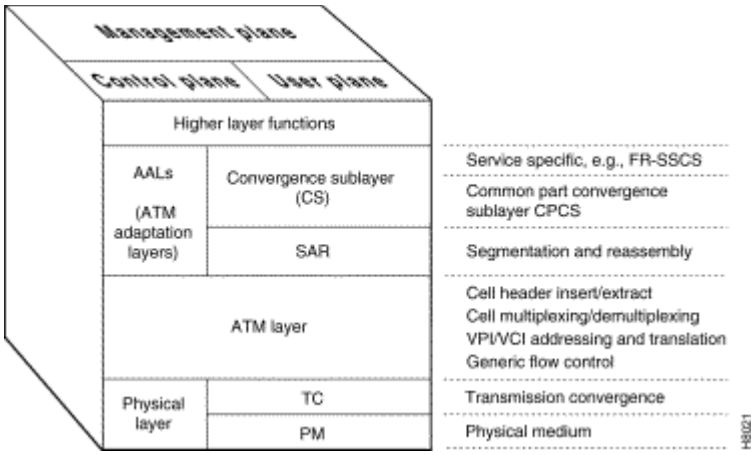
ATM is a CONNECTION-ORIENTED scheme, in which switches create a VIRTUAL CIRCUIT between the sender and receiver of a call that persists for the duration of the call. It is a PACKET SWITCHING system, which breaks down messages into very small, fixed length packets called CELLS generally 53 bytes in length (48 bytes of data plus a 5-byte header). The advantage conferred by such small cells is that they can be switched entirely in hardware, using custom chips, which makes ATM switches very fast (and potentially very cheap).

The ASYNCHRONOUS part of the name refers to the fact that although ATM transmits a continuous stream of cells, some cells may be left empty if no data is ready for them so that precise timings are not relevant. This is ATM's greatest strength, as it enables flexible management of the QUALITY OF SERVICE so; an operator can offer different guaranteed service levels (at different prices) to different customers even over the same line. This ability will enable companies to rent VIRTUAL PRIVATE NETWORKS based on ATM that behave like private leased lines but in reality share lines with other users.

- **Layers** of ATM model (AAL, ATM Layer, ATM Physical Layer)



Downloaded from CSIT Tutor



AAL (ATM Adaptation Layer): A software layer that accepts user data, such as digitized voice, video or computer data, and converts to and from cells for transmission over an ASYNCHRONOUS TRANSFER MODE network. AAL software mostly runs at the end-points of a connection, though in a few circumstances AAL software is run inside an ATM switch. AAL includes facilities to carry traffic that uses other network protocols, such as TCP/IP, over ATM.

Frame Relay

Frame relay has evolved from X.25 packet switching and objective is to reduce network delays, protocol overheads and equipment cost. Error correction is done on an end-to-end basis rather than a link -to-link basis *as* in X.25 switching. Frame relay can support multiple users over the same line and can establish a permanent virtual circuit or a switched virtual circuit.

Frame relay is considered to be a protocol, which must be carried over a physical link. While useful for connection of LANs, the combination of low throughput, delay variation and frame discard when the link is congested will limit its usefulness to multimedia.

Packet switching was developed when the long distance digital communication showed a large error rate.

- To reduce the error rate, additional coding bits were introduced in each packet in order to introduce redundancy to detect and recover errors.
- But in the modem high speed telecommunication a system, this overhead is unnecessary and infect counterproductive.
- Frame relay was developed for taking the advantage of the high data rates and low error rates in the modem communication system.
- The original packet switching networks were designed with a data rate at the user end of about 64 kbps.
- But the frame relay networks are designed to operate efficiently at the user's data rates up to 2 Mbps.
- This is possible practically because most of the overhead (additional bits) are striped off.
- Frame relay is a virtual circuit wide area network which was designed in early 1990s.
- Frame relay also is meant for more efficient transmission scheme than the X.25 protocol.
- Frame Relay is used mostly to route Local Area Network protocols such *as* IPX or *TCP/IP.*

Downloaded from CSIT Tutor

- The biggest difference between Frame Relay and X.25 is that X.25 guarantees data integrity and network managed flow control at the cost of some network delays. Frame Relay switches packets end-to-end much faster, but there is no guarantee of data integrity at all.

Features of frame relay:

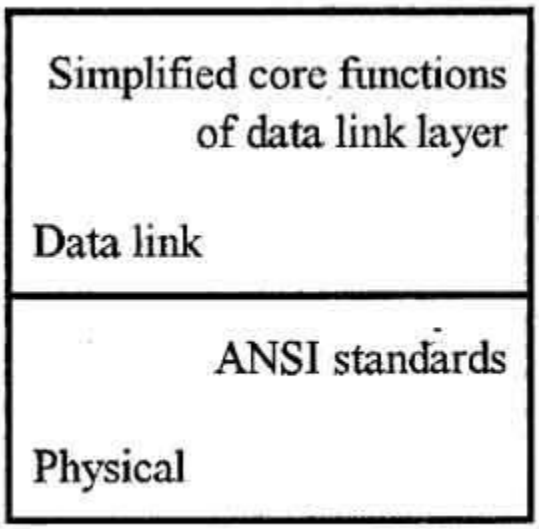
i. Frame relay operates at a high speed (1.544 Mbps to 44.376 Mbps).
ii. Frame relay operates only in the physical and data link layers. So it can be easily used in Internet.
iii. It allows the bursty data.
iv. It has a large frame size of 9000 bytes. So it can accommodate all local area network frame sizes.
v. Frame relay can only detect errors (at the data link layer). But there is no flow control or error control.
vi. The damaged frame is simply dropped. There is no retransmission. This is to increase the speed. So frame relay needs a reliable medium and protocols having flow and error control.

Frame Format

- The DLCI length is 10 bits
- There are two EA locations. The value of the first one is fixed at 0 and the second at
- 1 is set in the DE (Discard Eligibility) for the part that can be discarded first when congestion occurs
- The data size may vary up to 4096 bytes.

Frame relay layers

• Frame relay has only two layers i.e. physical layer and data link layer.



Frame relay layers

Physical layer

• Frame relay supports ANSI standards.

Downloaded from CSIT Tutor

• No specific protocol is defined for the physical layer. The user can use any protocol which is recognized by ANSI.

Data link layer

• A simplified version of HDLC is employed by the frame relay at the data link layer.

• A simpler version is used because flow control and error correction is not needed in frame relay.

Downloaded from CSIT Tutor

2.2 Multimedia Networking

• New applications

In recent years, there has been an explosive growth of new applications on the Internet like streaming video, IP telephony, teleconferencing, interactive games, virtual world, distance learning, and so on. Those multimedia networking applications are referred as continuous-media applications and require services different from those for traditional elastic applications like e-mail, Web, remote login, etc. They are also different from download-and-then-play applications. Especially, the new applications require high quality on the communication latency and the latency variation (delay-sensitive) but may not require high quality on the error rate (loss-tolerant). One key issue for supporting new multimedia networking applications is how to get the high quality for the communication latency on the best-effort Internet which provides no latency guarantee. Another key issue is how to improve the Internet architecture to provide support for the service required by multimedia applications.

• Examples of new applications

Streaming stored audio and video.

Applications have the following key features:

– Stored media, the contents has been prerecorded and is stored at the server. So, a user may pause, rewind, or fast-forward the multimedia contents. The response time to the above actions should be in the order of 1-10 seconds.

– Streaming, a user starts playout a few seconds after it begins receiving the file from the server. So, a user plays out the audio/video from one location in the file while it is receiving later parts of the file from the server. This technique is called streaming and avoids having download the entire file before starting playout.

– Continuous playout, once playout begins, it should proceed based on the original timing of the recording. This requires high quality on the end-to-end delay.

Streaming live audio and video.

Applications are similar to traditional radio and television, except that audio/video contents are transmitted on the Internet. In these applications, many clients may receive the same program. A key issue here is how to deliver the program efficiently to multiple clients on the Internet. IP multicasting technologies play a key role for this. Similar to streaming stored audio and video applications, applications here require continuous playout and high quality on the end-to-end delay.

Real time interactive audio and video.

Applications allow users using audio/video to communicate with each other in real time. Real-time interactive audio on the Internet is known as Internet phone. Applications in this category require very high quality on the end-to-end delay, usually a fraction of one second.

Downloaded from CSIT Tutor

• Hurdles for multimedia in today's Internet

The Internet Protocol (IP) used in the Internet provides connectionless best effort service for transmitting datagrams. The IP does not guarantee the end-to-end delay nor the uniform delay for all datagrams in a same packet stream. The variations of packet delays within the same packet stream is called packet jitter. The end-to-end delay and packet jitter in the Internet are major hurdles for multimedia applications on the Internet.

• How to overcome hurdles

There are many approaches discussed for overcoming the hurdles mentioned above. At one extreme, it is argued that fundamental changes to the Internet should be made so that the users can explicitly reserve the bandwidth on every link in the path for transmitting the packets.

On the other hand, it is argued that fundamental changes are difficult and incremental improvements over the best-effort IP are more practical. Especially, the improvements include:

– The Internet Service Providers (ISP) should scale/upgrade their networks well to meet the demands. The upgrade includes more bandwidth and caches in networks for heavily accessed data.

– Content distribution networks (CDNs), replicate stored contents and put the contents at edges of the Internet.

– Multicast overlay networks for sending data to a huge number of users simultaneously.

Another approach is differentiated services (Diffserv). In this approach, small changes at the network and transport layers are required and scheduling/policing schemes are introduced at edges of the network. The idea is to introduce traffic classes, assign each datagram to one of the classes, and give datagrams different levels of services based on their class.

Streaming stored audio and video

• Overview

In these applications, clients request audio/video data stored at servers. Upon client's request, servers send the data into a socket connection for transmission. Both TCP and UDP socket connections have been used in practice. The data are segmented and the segments are encapsulated with special headers appropriate for audio/video traffic. The real time protocol (RTP, will be discussed later) is a public-domain standard for encapsulating such segments.

Audio/video streaming applications usually provide user interactivity which requires a protocol for client/server interaction. The real time streaming protocol (RTSP) is a public-domain protocol for this purpose.

Clients often request data through a Web browser. A separate helper application (called media player) is required for playing out the audio/video. Well used helpers include RealPlayer and MediaPlayer.

• Access audio/video through Web server

The stored audio/video files can be delivered by a Web server or by an audio/video streaming server. When an audio file is delivered by a Web server, the file is treated as an ordinary object in the server's file system, like HTML and JEPG files. To get the file, a client establishes a TCP

Downloaded from CSIT Tutor

connection with the server and sends an HTTP request for the object. On receiving the request, the Web server encapsulates the audio file in an HTTP response message and sends the message back to the TCP connection. It is more complicated for the video case because usually the sounds (audio) and images are stored in two different files. In this case, a client sends two HTTP requests over two separate TCP connections and the server sends two responses, one for sounds and the other for images, to the client in parallel. It is up to the client to synchronize the two streams.

• Sending multimedia from a streaming server to a helper application

Audio/video files can be delivered by a streaming server to a media player. Streaming servers include those marketed by RealNetworks and Microsoft, and those of public-domain servers. With a streaming server, audio/video files can be transmitted over UDP which has much smaller end-to-end delay than TCP.

• Real-Time Streaming Protocol (RTSP)

RTSP is a protocol which allows a media player to control the transmission of a media stream. The control actions include pause/resume, repositioning of playback, fast-forward, and rewind. RTSP messages use a different port number from that used in the media stream and can be transmitted on UDP or TCP.

Making the best of the best-effort service

• Limitation of best-effort service

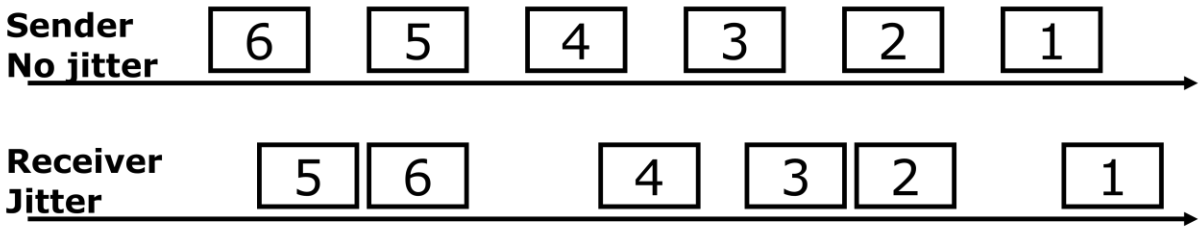
Packet loss, IP provides the best-effort service but does not guarantee the delivery of packets. Packets may be discarded due to congestions.

End-to-end delay, IP does not guarantee the end-to-end delay either. The time for transmitting a packet may vary due to the conditions of the network. Also, in order to guarantee the delivery, positive acknowledgement and retransmission are used in TCP. The cost for realizing the reliable transmission in TCP is a longer end-to-end delay.

Packet jitter, since the end-to-end delay for each packet may depend on the conditions of the network, the delays of packets in the same packet stream may vary. Especially, the packets may arrive to the receiver in a wrong order.

• Removing jitter at the receiver for audio

Downloaded from CSIT Tutor



In applications like Internet phone or audio-on-demand, it is up to the receiver to remove the jitters. Common techniques used include sequence number, timestamp, and delaying playout. The sender can put a sequence number on every packet sent and the receiver can use the sequence number to recover the correct order of the received packets. Timestamp is similar to sequence number, the sender stamps each packet with the time at which the packet is generated. In order to get the correct order from the sequence number and timestamp for a sequence of packets, the receiver need to receive all of the packets in the sequence. Playout delay is used for this purpose. The playout delay should be long enough to receive all packets in a subsequence of packets which can be played. On the other hand, the delay should be short enough so that the user will not notice the delay. The playout delay can be either fixed or adaptive.

Fixed playout delay, the receiver plays out each packet exactly q msecs after the packet is generated. Usually, q is up to a few hundreds msecs.

Adaptive playout delay, the receiver estimate the network delay and the variance of the network delay at the beginning of each talk, and adjusts the playout delay accordingly.

• Recovering from packet loss

A major scheme for handling packet loss for elastic applications is retransmission. However, this scheme does not work well for applications with strict end-to-end delay constraint. Internet phone applications usually use loss anticipation schemes to handle packet loss.

Forward error correction (FEC) is one of such schemes. The basic idea of this scheme is to include redundant information in the original packet stream. The redundant information can be used to reconstruct the lost packet. One approach for the FEC scheme is to send the exclusive OR of every n packets as a redundant packet. If any one of the n + 1 packet is lost, the receiver can reconstruct it. However the scheme does not work if two or more of the n + 1 packets are lost. Another approach is to send two copies of the same packet, usually one is the original

Downloaded from CSIT Tutor

packet and the other is a short version (lower-resolution audio) of the packet. An example is that the short version of packet i is sent together with packet i + 1. FEC uses extra bandwidth of networks.

Interleaving is another loss anticipation scheme. This scheme resequences units of audio data before transmission so that the original adjacent units are separated by some distance in the transmitted stream. The receiver rearranges the received stream into its original order before it is resequenced. If a transmitted packet is lost, only a small fraction of each original packet is lost and the quality of the voice may not be damaged much. Interleaving does not use extra bandwidth but introduces extra end-to-end delay.

Receiver-based repair of damaged audio stream. This scheme reconstructs a lost packet using the other received packets based on the fact that there are large amount of short term self-similar signals in audio data, especially for speech. A simplest approach is packet repetition, using the immediate previous packet to replace the lost one. Another approach is interpolation, using the packets before and after the loss to interpolate a packet to cover the loss.

Protocols for real-time interactive applications

• Real Time Protocol (RTP)

In multimedia applications, header fields are appended to audio/video packets for transmission. These header fields include sequence number and timestamps. RTP is a standard for the packet structure which includes the fields for audio/video data, sequence number, timestamp, and other fields.

Usually, the media data is encapsulated in RTP packets which are encapsulated in UDP segments.

RTP packet header fields include

RTP Header

– payload type, 7 bits, used to indicate the type of encoding for audio and video;
– sequence number, 16 bits, incremented by one for each RTP packet sent;
– timestamp, 32 bits, used to give the sampling instant of the first byte in the RTP packet data;
– synchronization source identifier (SSRC), 32 bits, used to identify the source of the RTP stream;
– miscellaneous fields

• RTP control protocol (RTCP)

RTCP is a protocol that a networked multimedia application can use in conjunction with RTP. RTCP packets do not carry audio/video data but contain sender/receiver reports which include the statistics on the number of RTP packets sent, number of packets lost, and interarrival jitters. RTCP packets are sent periodically. There are two types of RTCP packets.

The RTCP packets used by receiver include

Downloaded from CSIT Tutor

– the SSRC of the RTP stream for which the reception report is generated;
– the fraction of the packets lost within the RTP stream;
– the last sequence number received in the stream of RTP packets; and
– the interarrival jitter.
The RTCP packets used by sender include
– the SSRC of the RTP stream;
– the timestamp and real time of the most recently generated RTP packet in the stream;
– the number of packets sent in the stream; and
– the number of bytes sent in the stream.
In an RTP session, if the number of receivers is large, the overhead of the RTCP packets generated by receivers can be large. The period of sending RTCP packets for receivers in a large multicast tree should be carefully designed to limit the overhead.

• Session initiation protocol (SIP)

SIP provides mechanisms for the following.
– It establishes calls between a caller and a callee over an IP network. It allows the caller to notify the callee that it wants to start a call. It allows the participants to agree on media encodings and to end a call.
– It allows the caller to determine the current IP address of the callee. Users may have multiple or dynamic IP addresses.
– For call management like adding new media streams, changing the encoding, inviting new participants, call transfer, and call holding.
Setting up a call to a known IP address. The caller initiates the call by sending an INVITE message which includes an identifier for the callee (e.g., the name of callee and his/her IP address), an identifier for the caller, the encoding scheme used in the audio data, the port number through which the caller wants to receive the RTP packets. The callee, after receiving the INVITE message, sends an SIP response message which includes an OK message, an indication of callee's IP address, desired encoding scheme for reception, and port number for the conversation. After receiving the SIP response message, the caller sends the callee an ACK message. After that, the call connection is set-up. The SIP messages are transmitted through a well known port number 5060 for SIP. The RTP packets are transmitted on different connections. SIP addresses can be in the form of IP addresses or that similar to email addresses.
Name translation and user location. SIP proxy and SIP registrar are used to track the users with multiple or dynamic IP addresses.

• H.323

H.323 is popular standard for real-time audio and video conferencing among end systems in the Internet. The standard includes the following:
– A specification for how endpoints negotiate common audio/video encodings.
– H.323 mandates RTP for audio and video data encapsulation and transmission over the network.

Downloaded from CSIT Tutor

– A specification for how endpoints communicate with their respective gatekeepers (a device similar to and SIP registrar).
– A specification for how Internet phones communicate through a gateway with ordinary phones in the public circuit-switched telephone networks.

Beyond best effort

• We have discussed a number of techniques such as sequence numbers, timestamps, FEC, RTP, and H.323 for improving the performances of the best-effort Internet for multimedia applications which require high quality of service on the end-to-end delay. However, these techniques do not change the best-effort nature of the Internet. Now we discuss some technologies which are used to provide the true quality of service for multimedia applications. The central idea for those technologies is to add new architectural components to the Internet to change its best-effort nature. Those technologies have been under active discussion in the Internet Engineering Task Force (IETF) working groups for Diffserv, Intserv, and RSVP.

To see how to change the best-effort nature of the Internet by adding new architectural components, we start from a simple example. Assume that routers R1 and R2 are gateways for networks N1 and N2, respectively, and are connected by a link of limited bandwidth. There are two traffic streams, one is a multimedia application and the other is an FTP from N1 to N2. By the best-effort IP routing, the datagrams for both streams will be transmitted without any priority by routers. In this case, the busty arrival of FTP packets may delayed the transmission of multimedia packets although the multimedia application has strict requirement on the end-to-end delay while the FTP does not. One way to solve this problem seems to add new functions to IP such that the

protocol can treat the two streams differently. There are a few principles for this.
– First, classification of packets is needed to allow a router to distinguish among the packets belonging to different classes of traffic.
– It is desirable to provide a degree of isolation among traffic flows so that one flow is not adversely affected by another misbehavior flow.
– While providing isolation among flows, it is desirable to use resources (like bandwidth of links and buffers) as efficiently as possible.
– If resources are not enough, a call admission process is needed in which flows declare their QoS requirements and are then either admitted to the network (at the required QoS) or blocked from the network (if the required QoS can not be provided by the network).
Scheduling and policing mechanisms are used to provide QoS guarantees.

• Scheduling

A scheduling mechanism is a scheme for selecting packets for transmission from an output link queue with packets of multiple data streams.
A simple scheduling scheme is first-in-first-out (FIFO). In FIFO scheduling, packets are transmitted in the order of their arrival.

Downloaded from CSIT Tutor

Priority queuing is another scheduling scheme. In this scheme, arrived packets are classified into priority classes at the output queue. The priority of a packet may depend on an explicit marking in its header, its source/destination addresses, port numbers, or other criteria. Usually, each priority class has its own queue. The priority queue scheduling chooses a packet by FIFO scheme for transmission from the highest priority class that has a non-empty queue.

In round robin scheduling, packets of each data stream are put in a distinct queue and the queues are served in a circular manner. Assume that there are n queues. In the simplest form, the scheduling scheme serves queue 1 (selects a packet by FIFO for transmission from queue 1 if queue 1 is not empty), then serves queue 2. In general, queue i + 1 mod (n + 1) is served after queue i is served.

Weighted fair queuing (WFQ) is a generalization of round robin. In this scheme, packets are classified and each class is given a distinct queue. Each queue i is also assigned a weight w_i. Queues are served in a round robin manner with queue i be guaranteed to receive a fraction of service equal to $w_i/(sum_j (w_j))$, where the sum_j is taken over all classes that have non-empty queues.

• Policing

Policing schemes are used to specify the data rate at which a data stream is allowed to enter a network. Criteria for policing include:

– Average rate, number of packets per time interval at which packets of a data flow can be sent into the network. A key issue is the interval of time over which the average rate will be policed.

– Peak rate, the maximum number of packets that can be sent into the network over a short period of time.

– Burst size, the maximum number of packets that can be sent into the network over an extremely short interval of time.

Leaky bucket is a mechanism used to specify policing limits shown above. A leaky bucket consists of a bucket of size b. When the bucket is full, it has b tokens. If the bucket has less than b tokens, new tokens are added to the bucket with a constant rate of r tokens per second until the bucket becomes full. In the leaky bucket policing, when a packet is transmitted into a network, it must first remove a token from the bucket. If the bucket is empty, the packet must wait for a token. The burst size defined by the leaky bucket is b, the peak rate over time t is rt + b, and the average rate is r.

The leaky bucket can be combined with the weighted fair queue. Each queue i is associated with a leaky bucket with size b_i and new token generation rate r_i.

• Integrated service and differentiated service

The principles and mechanisms discussed above are used in two architectures, integrated service (Intserv) and differentiated service (Diffserv), proposed to providing QoS in the Internet. Intserv is a framework developed within the IETF to provide individualized QoS guarantees to individual application sessions. Diffserv provides the ability to handle different classes of traffics in different ways within the Internet.

• Intserv

Downloaded from CSIT Tutor

There are two key features in the Intserv architecture.
– Reserved resources, a router is required to know what amounts of resources have been reserved for ongoing sessions.
– Call setup, a session which requires a QoS guarantee must reserve resources at every router on the path for transmission. The session must send the traffic characterization and specification of the desired QoS to routers. Routers determine if the call of the session is admitted or not. Routers reserve the resources to guarantee the QoS required if they decide to admit the call.
The RSVP protocol is used for the call setup.

• Diffserv

Intserv provides the QoS guarantee for each individual session. This has advantages but also introduces problems. Especially, the per-flow resource reservation may give significant workload to routers. Also Intserv does not allow for more qualitative or relative definitions of service distinctions.
The Diffserv architecture is proposed to provide scalable and flexible service. In Diffserv, different classes of traffics can be treated in different ways in the Internet. There are two sets of functional elements in the Differ architecture.
– Edge functions: packet classification and traffic conditioning, packets arriving to the edge router are first classified based on the values of one or more packet header fields.
– Core function: forwarding, a classified packet is forwarded to its next hop router according to the per-hop behavior associated with the packet's class. The per-hop behavior affects how a router's buffer and link bandwidth are shared with other classes of traffics.

• Resource Reservation Protocol (RSVP)

The resource reservation protocol (RSVP) is used by application sessions to reserve resources in the Internet. Especially, RSVP is used to reserve bandwidth for multicast trees (unicast is treated as a degenerate case of multicast). RSVP is receiver-oriented. The receiver of a data flow initiates and maintains the resource reservation. RSVP is a protocol used by sessions to send the reservation request and does not specify how the network provides the reserved resources. It is not a routing protocol either and does not determine the links in which the reservations are to be made. RSVP operates in a two-pass manner. A transmitting source advertises its content by sending an RSVP path message through a multicast tree, indicating the bandwidth required for the content, the timeout interval, and information about the upstream path to the source. Each receiver sends an RSVP reservation message upstream on the multicast tree. The reservation message specifies the rate at which the receiver wants to receive the data. When a router receives a reservation message, it first checks if its downstream links can accommodate the reservation. If yes, it adjusts its packet scheduler to accommodate the reservation and sends a reservation upstream on the multicast tree. Otherwise it rejects the reservation and sends an error message to the corresponding receivers.

Downloaded from CSIT Tutor

2.3 Network Management

In the early days, network was small and local. Network manager's job includes:

- Installation: attach PCs, printers, etc. to LAN
- Configuration: NICs, protocol stack, user app's shared printers, etc.
- Testing: Ping was sufficient to "manage" network
- More devices: bridge, router

Job was manageable. Above tasks only deal with configuration. Ongoing maintenance issues are:

- How to optimize performance?
- How to handle failures and network changes?
- How to extend network capacity?
- How to account for network usages?
- How to solve network security issues?

In the past, the network manager might take all the responsibilities. Today the task has divided into specialties:

- Server admin
- System admin
- Network admin
- Security specialist
- Different certifications for these
    - Cisco, Novell, Microsoft, Sun, $(ISC)^2$ etc.

Today, networks are larger and more complicated, so more demands on network manager.

- How to monitor and control the network effectively and timely?
- Management tools are needed

Network-based management tools: use the network to manage the network (remotely)

- To control
    - Simple Network Management Protocol (SNMP)
    - Management Information Base (MIB)
    - Network Management System (NMS)
- To monitor
    - Remote Monitor (RMON1)

Definition by Saydam (in Journal of Networks and System Management, published in Dec. 1996):

"Network management includes the deployment, integration and coordination of the hardware, software, and human elements to monitor, test, poll, configure, analyze, evaluate, and control the network and element resources to meet the real-time, operational performance, and Quality of Service requirements at a reasonable cost."

In brief:

- Network management is mostly a combination of local and remote configuration and management with software.
- Remote network management is accomplished when one computer is used to monitor, access, and control the configuration of other devices on the network.

Benefits of Network Management

- Detecting failure of an interface card at a host or a router
- Host monitoring

Downloaded from CSIT Tutor

- Monitoring traffic to aid in resource deployment
- Detecting rapid changes in routing tables
- Monitoring for SLAs (Service Level Agreements)
- Intrusion detection

ISO Network Management Categories

- Performance Management
- Fault Management
- Configuration Management
- Security Management
- Accounting Management

Performance Management

- Concerned with
    - Response time
    - Utilization
    - Error rates, etc.
- Must collect and analyze data
    - Number and type of packets
    - Might also rely on simulations

Fault Management

- Preventions, detection and isolation of abnormal behavior
    - May be caused by malfunction, cable issue, the janitor, etc.
- Traffic, trends, connectivity, etc.
    - SNMP **polls**
    - **Alarms** for automatic fault detection
    - Monitor statistics
    - Timeliness etc.

Configuration Management

- Device configuration
    - May be done locally or remotely
- Network configuration
    - Sometimes called “capacity mgmt”
    - Critical to have sufficient capacity
- Desirable to automate as much as possible
    - For example, DHCP and DNS
- Extensions to SNMP MIB

Security Management

- Control access to network/resources
    - Authentication: who goes there?
    - Authorization: are you allowed to do that?
    - Firewalls

Downloaded from CSIT Tutor

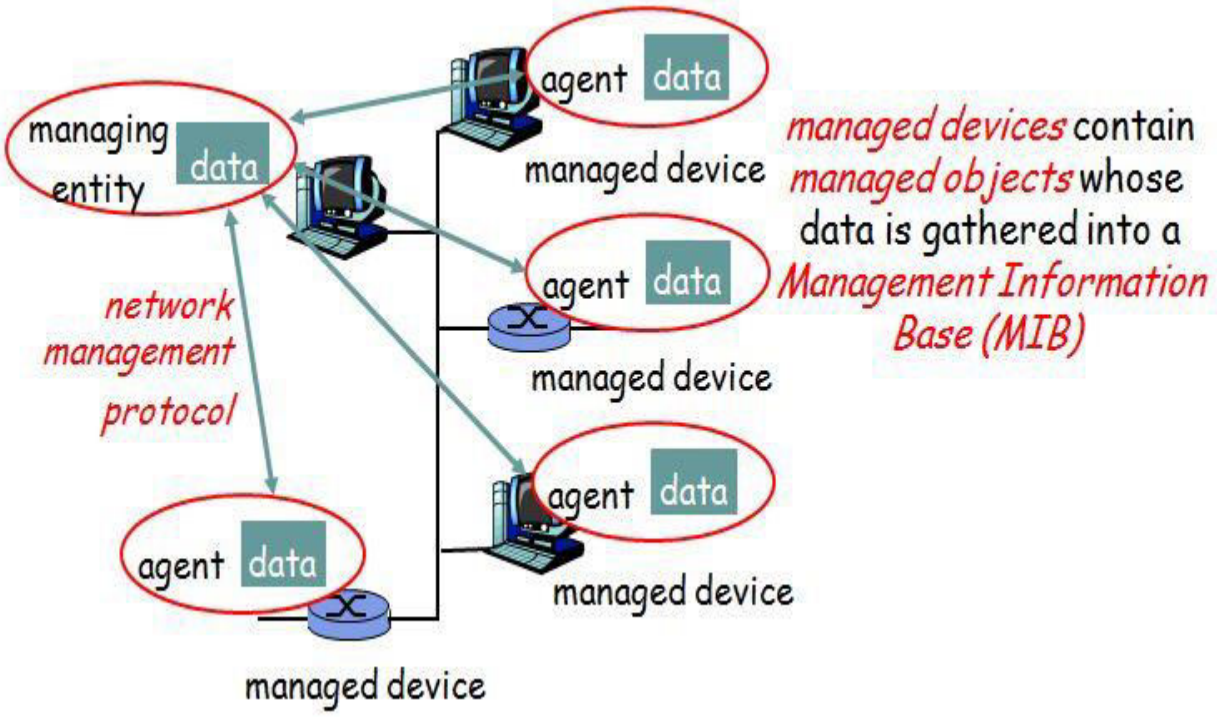
- Intrusion detection systems (IDS)
- Notification of (attempted) breaches, etc.

- Critical to always authenticate participants
- SNMPv1 has very little security
- SNMPv3 has lots of security built in

Accounting Management

- Measuring the usage of network resources in order to distribute costs and resources
- Allows the network manager to specify, log and control user and device access to network resources
- E.g., monitoring the use of a server by users in a specific department and charging the department accordingly

Infrastructure for Network management



- Managed Device
  - Devices to be monitored/controlled, e.g., router, switch, hub, bridge, workstation.
  - A managed device may have several managed objects to be managed
  - Managed objects mean pieces of hardware, and sets of configuration parameters for hardware and software such as routing protocols
  - A software (agent) is installed to provide **access** to information/parameters (data) about the device, which is called Management Information Base (MIB)
- Managing Entity
  - An application used by the manager/Admin to do network management
  - It controls the collection, processing, analysis, and/or display of network management information
  - PC, notebook, terminal, etc., installed with a software called Network Management System (NMS)
  - NMS displays/analyzes data from management agents

Downloaded from CSIT Tutor

- Network Management Protocol
    - Runs between the managing entity and the managed devices
    - The managing entity can query the status of the managed devices and take actions at the devices via its agents
    - Agents can use the protocol to inform the managing entity of exceptional events
    - E.g., SNMP: Simple Network Management Protocol
- Managing agents located at managed devices are periodically queried by the managing entity through a network management protocol.

Internet-Standard Management Framework

This addresses

- What is being monitored? And what form of control can be exercised by the network administrator?
- What is the specific form of the information that will be reported and/or exchanged?
- What is the communication protocol for exchanging this information?

4 parts

- MIB (Management Information Base)
- SMI (Structure of Management Information)
- SNMP (Simple Network Management Protocol)
- Security and administration capabilities

MIB

- Represented as a collection of managed objects that together form a virtual information store
- Might be a counter, such as the number of IP datagrams discarded at a router due to errors in the IP datagram header; or the number of carrier sense errors in an Ethernet interface card; descriptive information such as the version of software running on a DNS server, status information such as whether a particular device is functioning properly; or protocol-specific information such as a routing path to a destination
- MIB objects thus define the management information maintained by a managed device
- Related MIB objects are gathered into MIB modules

SMI

- Data definition language
- Defines the data types, an object model, and rules for writing and revising management information
- MIB objects are specified in this data definition language

SNMP

- Protocol used for conveying information and commands between a managing entity and an agent executing on behalf of that entity within a managed network device

Security and administration capabilities

- The addition of these capabilities represents the major enhancement in SNMPv3 over SNMPv2

Downloaded from CSIT Tutor

Network management example

- To get value of MIB variable from mgmt agent
    1. Mgmt app (part of NMS) on managing entity passes request to mgmt process
    2. Mgmt process calls network mgmt protocol (e.g., SNMP)
    3. SNMP constructs Get-Request packet and sent it to the managed device through the network
    4. Mgmt agent on managed device receives Get-Request
    5. Agent process accesses requested value
    6. SNMP constructs Get-Response packet and sent it to managing entity through the network
    7. Mgmt process on managing entity receives response
    8. Mgmt process passes data to mgmt app

Network Management Overhead

- There is overhead in terms of
    - CPU cycles to generate and process information/packets
        - May require dedicated Managing Entity
    - Bandwidth usage for sending request and receiving responses
- A tradeoff between cost and benefit

Additional Network Management Capabilities

- For efficiency, multiple values can be constructed in a single Get-Response packet
- Can traverse MIB in logical order
- Mgmt agent can send unsolicited messages
    - These are known as **traps**
    - E.g., if a device goes down
- Can request info from probes or remote monitors (RMON)
    - Monitoring activity (traffic) on a network segment

Downloaded from CSIT Tutor