UDP/TCP

* The data link layer is responsible for delivery of frames between two neighbouring nodes over a link.
  + This is called node-to-node delivery.
* The network layer is responsible for delivery of datagrams between two hosts.
  + This is called host-to-host delivery.
* Real communication takes place between two processes (application programs). The transport layer is responsible for process-to-process delivery.
  + The delivery of a packet, part of a message, from one process to another.

At any moment, several processes may be running on the source host and several on the destination host. To complete the delivery, we need a mechanism to deliver data from one of these processes running on the source host to the corresponding process running on the destination host.

The most common ways to achieve process-to-process communication is through the client/server paradigm. A process on the local host (i.e. client) needs services from a process usually on the remote host (i.e. server).

A remote computer can run several server programs at the same time, just as local computers can run one or more client programs at the same time.

For communication, we must define the following:

* Local host
* Local process
* Remote host
* Remote process

Whenever we need to deliver something to one specific destination among many, we need an address.

* At the data link layer, we need a MAC address to choose one node among several nodes if the connection is not point-to-point.
  + A frame in the data link layer needs a destination MAC address for delivery and
  + A source address for the next node's reply.
* At the network layer, we need an IP address to choose one host among millions.
* At the transport layer, we need a transport layer address, called a port number, to choose among multiple processes running on the destination host.
  + The destination port number is needed for delivery;
  + The source port number is needed for the reply.

In the Internet model, the port numbers are 16-bit integers between 0 and 65,535.

The client program defines itself with a port number, chosen randomly by the transport layer software running on the client host. However, server port number, cannot be chosen randomly.

The Internet has decided to use universal port numbers for servers; these are called well-known port numbers.

Every client process knows the well-known port number of the corresponding server process.

The lANA (Internet Assigned Number Authority) has divided the port numbers into three ranges:

* Well-known ports. The ports ranging from 0 to 1023 are assigned and controlled by lANA.
* Registered ports. The ports ranging from 1024 to 49,151 are not assigned or controlled by lANA.
  + They can only be registered (not controlled) with lANA to prevent duplication.
* Dynamic ports. The ports ranging from 49,152 to 65,535 are neither controlled nor registered.
  + They can be used by any process. These are the ephemeral ports.

Socket address: The combination of an IP address and a port number is called a socket address.

* The client socket address defines the client process uniquely
* The server socket address defines the server process uniquely
* A transport layer protocol needs a pair of socket addresses:
  + the client socket address and
  + The server socket address.
* These four pieces of information are part of the IP header and the transport layer protocol header.
  + The IP header contains the IP addresses;
  + The UDP or TCP header contains the port numbers.

UDP

The User Datagram Protocol (UDP) is called a connectionless, unreliable transport protocol.

UDP is a very simple protocol using a minimum of overhead.

If a process wants to send a small message and does not care much about reliability, it can use UDP.

Sending a small message by using UDP takes much less interaction between the sender and receiver.

User Datagram

UDP packets, called user datagrams, have a fixed-size header of 8 bytes.

Format of a user datagram.

* Source port number. It is 16 bits long.
  + If the source host is the client the port number is requested by the process and chosen by the UDP software running on the source host.
* Destination port number. It is also 16 bits long.
* Length. This is a 16-bit field that defines the total length of the user datagram, header plus data.
* Checksum. This field is used to detect errors over the entire user datagram

UDP Operation

* Connectionless service : each user datagram sent by UDP is an independent datagram.
  + The user datagrams are not numbered.
  + There is no connection establishment and no connection termination, each request must be small enough to fit into one user datagram.
  + Only those processes sending short messages should use UDP.
* Flow and Error Control
  + There is no flow and error control mechanism in UDP except for the checksum.
  + When the receiver detects an error through the checksum, the user datagram is silently discarded.
  + Process using UDP should provide Flow and Error Control mechanisms.
* Encapsulation and Decapsulation : To send a message from one process to another, the UDP protocol encapsulates and decapsulates messages in an IP datagram.

TCP

TCP Services:

TCP is called a connection-oriented, reliable transport protocol. It adds connection-oriented and reliability features to the services of IP.

Process-to-Process Communication : TCP provides process-to-process communication using port numbers.

Stream Delivery Service

TCP, is a stream-oriented protocol.

TCP, allows the sending process to deliver data as a stream of bytes and allows the receiving process to obtain data as a stream of bytes.

Sending and Receiving Buffers: Because the sending and the receiving processes may not write or read data at the same speed, TCP needs buffers for storage.

Full-Duplex Communication : TCP offers full-duplex service, in which data can flow in both directions at the same time.

Connection-Oriented Service

TCP is a connection-oriented protocol. When a process at site A wants to

send and receive data from another process at site B, the following occurs:

1. The two TCPs establish a connection between them.

2. Data are exchanged in both directions.

3. The connection is terminated.

Reliable Service

TCP is a reliable transport protocol. It uses an acknowledgment mechanism to check the safe and sound arrival of data.

**Flow Control:**

TCP uses a sliding window to handle flow control.

**Error Control:**

Error detection and correction in TCP is achieved through the use of three simple tools: checksum, acknowledgment, and time-out.

checksum

Each segment includes a checksum field which is used to check for a corrupted segment. If the segment is corrupted, it is discarded by the destination TCP and is considered as lost. TCP uses a 16-bit checksum that is mandatory in every segment.

Acknowledgment

TCP uses acknowledgments to confirm the receipt of data segments.

**Congestion**

Congestion in a network occurs if the load on the network is greater than the capacity of the network.

Congestion in a network or internetwork occurs because routers and switches have queues (buffers) that hold the packets before and after processing.

When a packet arrives at the incoming interface, it undergoes three steps before departing,

The packet is put at the end of the input queue while waiting to be checked

The processing module of the router removes the packet once it reaches the front of the queue and uses its routing table and the destination address to find the route.

The packet is put in the appropriate output queue and waits its turn to be sent.

two issues.

if the rate of packet arrival is higher than the packet processing rate, the input queues become longer and longer.

if the packet departure rate is less than the packet processing rate, the output queues become longer and longer.

when the load is much less than the capacity of the network, the delay is at a

minimum. This minimum delay is composed of propagation delay and processing delay,

both of which are negligible. However, when the load reaches the network capacity, the

delay increases sharply because we now need to add the waiting time in the queues (for all

routers in the path) to the total delay.

the delay becomes infinite when the load is greater than the capacity.

When a packet is delayed, the source, not receiving the acknowledgment,

retransmits the packet, which makes the delay, and the congestion, worse.

Throughput is the number of bits passing through a point in a second.

Extending this definition throughput in a network as the number of packets passing through the network in a unit of time.

When the load is below the capacity of the network, the throughput increases proportionally with the load.

but throughput declines sharply after the load reaches the capacity

When the load exceeds the capacity, the queues become full and the routers have to discard some packets. The sources retransmit the packets,

**TELNET**

TErminaL NETwork.

TELNET is a general-purpose client/server application program.

TELNET enables the establishment of a connection to a remote system in such a way

that the local terminal appears to be a terminal at the remote system.

When a user logs into a local timesharing system () and types at a terminal

the keystrokes are accepted by the terminal driver.

The terminal driver passes the characters to the operating system.

The operating system, in turn, interprets the combination of characters and

invokes the desired application program or utility.

When a user wants to access an application program located on a remote machine, she performs remote log-in.

Here the TELNET client and server programs come into use.

The user sends the keystrokes to the terminal driver, where the local

operating system accepts the characters but does not interpret them.

The characters are sent to the TELNET client, which transforms the characters to a universal character set

called network virtual terminal (NVT) characters and delivers them to the local TCP/IP protocol stack.

The commands or text, in NVT form, travel through the Internet and arrive at the TCP/IP stack at the remote machine.

Here the characters are delivered to the operating

system and passed to the TELNET server, which changes the characters to the

corresponding characters understandable by the remote computer.

However, the characters

cannot be passed directly to the operating system because the remote operating

system is not designed to receive characters from a TELNET server: It is designed to

receive characters from a terminal driver.

The solution is to add a piece of software called a pseudoterminal driver which pretends that the characters are coming from a terminal.

The operating system then passes the characters to the appropriate application program.

WWW

TheWWW today is a distributed clientJserver service, in which a client using a browser can access a service using a server. However, the service provided is distributed over many locations called sites

Each site holds one or more documents, referred to as Web pages.

Each Web page can contain a link to other pages in the same site or at other sites.

The pages can be retrieved and viewed by using browsers.

Client (Browser)

A variety of vendors offer commercial browsers that interpret and display a Web document, and all use nearly the same architecture.

Each browser usually consists of three parts: a controller, client protocol, and interpreters.

The controller receives input from the keyboard or the mouse and uses the client programs to access the document.

After the document has been accessed, the controller uses one of the interpreters to display the

document on the screen.

The client protocol can be one of the protocols described previously such as FfP or HTIP.

The interpreter can be HTML, Java, or JavaScript, depending on the type of document

Server

The Web page is stored at the server. Each time a client request arrives, the corresponding document is sent to the client.

Uniform Resource Locator

A client that wants to access a Web page needs the address. To facilitate the access of documents

distributed throughout the world, HTTP uses locators.

URL is a standard for specifying any kind of information on the Internet.

The URL defines four things: protocol, host computer, port, and path

Cookies

The World Wide Web was originally designed as a stateless entity. A client sends a

request; a server responds. Their relationship is over.

Today the Web has

other functions; some are listed here.

I. Some websites need to allow access to registered clients only.

2. Websites are being used as electronic stores that allow users to browse through the

store, select wanted items, put them in an electronic cart, and pay at the end with a

credit card.

3. Some websites are used as portals: the user selects the Web pages he wants to see.

4. Some websites are just advertising.

For these purposes, the cookie mechanism was devised.

Creation and Storage of Cookies

The creation and storage of cookies depend on the implementation; however, the principle

is the same.

854 CHAPTER 27 WWWAND HTTP

1. When a server receives a request from a client, it stores information about the client

in a file or a string. The information may include the domain name of the client, the

contents of the cookie (information the server has gathered about the client such as

name, registration number, and so on), a timestamp, and other information'depending

on the implementation.

2. The server includes the cookie in the response that it sends to the client.

3. When the client receives the response, the browser stores the cookie in the cookie

directory, which is sorted by the domain server name.

Using Cookies

The documents in the WWWcan be grouped into three broad categories: static, dynamic,

and active.

Static Documents

Static documents are fixed-content documents that are created and stored in a server.

The client can get only a copy of the document.

A Web page is made up of two parts: the head and the body.

The head contains the title of the page and other parameters that the browser will use.

The actual contents of a page are in the body, which includes the text and the tags.

Whereas the text is the actual infonnation contained in a page, the tags define the appearance of the document.

Another interesting tag category is the image tag. Nontextual information such as

digitized photos or graphic images is not a physical part of an HTML document. But

we can use an image tag to point to the file of a photo or image. The image tag defines

the address (URL) of the image to be retrieved. It also specifies how the image can be

inserted after retrievaL

A third interesting category is the hyperlink tag, which is needed to link documents together.

Dynamic Documents

A dynamic document is created by a Web server whenever a browser requests the document. When a request arrives, the Web server runs an application program or a script that creates the dynamic document. The server returns the output of the program or script as a response to the browser that requested the document.

Active Documents

For many applications, we need a program or a script to be run at the client site. These are called active documents.

One way to create an active document is to use Java applets.

HTTP

The Hypertext Transfer Protocol (HTTP) is a protocol used mainly to access data on the World Wide Web. HTTP functions as a combination of FTP and SMTP.

It is similar to FTP because it transfers files and uses the services of TCP.

However, it is much simpler than FTP because it uses only one TCP connection. There is no separate control connection;

Only data are transferred between the client and the server.

HTTP uses the services of TCP on well-known port 80.

HTTP is like SMTP

because the data transferred between the client and the server look like SMTP messages.

In addition, the format of the messages is controlled by MIME-like headers.

Unlike SMTP,

* The HTTP messages are not destined to be read by humans;
* They are read and interpreted by the HTTP server and HTTP client (browser).
* HTTP messages are delivered immediately.
  + SMTP messages are stored and forwarded
* The commands from the client to the server are embedded in a request message.
* The contents of the requested file or other information are embedded in a response message.

Although HTTP uses the services of TCP, HTTP itself is a stateless protocol.

The client initializes the transaction by sending a request message.

The server replies by sending a response.

A request message consists of a request line, a header, and sometimes a body.

A response message consists of a status line, a header, and sometimes a body.

The first line in a request message is called a request line;

The first line in the response message is called the status line.

Request type. This field is used in the request message. The request type is categorized into methods as GEY, PUT etc

URL. We discussed the URL earlier in the chapter.

Version. The most current version of HTTP is 1.1.

Status code. This field is used in the response message.

It consists of three digits.

* 100 range are only informational.
* 200 range indicate a successful request.
* 300 range redirect the client to another URL.
* 400 range indicate an error at the client site.
* 500 range indicate an error at the server site.

Header The header exchanges additional information between the client and the server.

For example, the client can request that the document be sent in a special format, or the server can send extra information about the document.

The header can consist of one or more header lines.

Each header line has a header name, a colon, a space, and a header value .

A header line belongs to one of four categories: general header, request header, response header, and entity header.

A request message can contain only general, request, and entity headers.

A response message, on the other hand, can contain only general, response, and entity headers.

Body The body can be present in a request or response message. Usually, it contains the document to be sent or received.