4.2 TRANSMISSION CONTROL PROTOCOL

Transmission Control Protocol (TCP) is a connection-oriented, reliable protocol. TCP defines connection establishment, data transfer, and connection teardown phases to provide a connection-oriented service.

TCP uses a combination of GBN and SR protocols to provide reliability. To achieve this ,TCP uses checksum(for error detection), retransmission of lost or corrupted packets, cumulative and selective acknowledgments, and timers.

TCP Services

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. TCP creates an environment in which the two processes seem to be connected by an imaginary "tube" that carries their bytes across the Internet. The sending process produces (writes to) the stream and the receiving process consumes (reads from) it.

Segments

At the transport layer, TCP groups a number of bytes together into a packet called a segment.

TCP adds a header to each segment (for control purposes) and delivers the segment to the network layer for transmission. The segments are encapsulated in an IP datagram and transmitted.

Full-Duplex Communication

TCP offers full-duplex service, where data can flow in both directions at the same time.

Multiplexing and Demultiplexing

Like UDP, TCP performs multiplexing at the sender and demultiplexing at the receiver.

Connection-Oriented Service

When a process at site A wants to send to and receive data from another process at site B, the following three phases occur:

The two TCP's establish a logical connection between them.

1.Data are exchanged in both directions.

2.The connection is terminated.

Reliable Service

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

TCP Features

The sequence number and the acknowledgment number is used in TCP.

Byte Number

TCP numbers all data bytes (octets) that are transmitted in a connection. Numbering is independent in each direction. When TCP receives bytes of data from a process, TCP stores them in the sending buffer and numbers them.TCP chooses an arbitrary number between 0 and $2^{32} - 1$ for the number of the first byte. For example, if the number happens to be 1057 and the total data to be sent is 6000 bytes, the bytes are numbered from 1057 to 7056.

Sequence Number

After the bytes have been numbered, TCP allot a sequence number to each segment that is being sent.The sequence number of the first segment is the ISN (initial sequence number), which is a random number.The sequence number of any other segment is the sequence number of the previous segment plus the number of bytes (real or imaginary) carried by the previous segment.

Acknowledgment Number

Communication in TCP is full duplex; when a connection is established, both parties can send and receive data at the same time. Each party numbers the bytes, usually with a different starting byte number.The sequence number in each direction shows the number of the first byte carried by the segment.Each party uses an acknowledgment number to confirm the bytes it has received.

TCP Segment

A packet in TCP is called a segment. The format of a segment is shown in Figure 4.2.1.

The segment consists of a header of 20 to 60 bytes, followed by data from the application program. The header is 20 bytes if there are no options and up to 60 bytes if it contains options.

Source port address. This is a 16-bit field that defines the port number of the application program in the host that is sending the segment.

Destination port address. This is a 16-bit field that defines the port number of the application program in the host that is receiving the segment.

Sequence number. This 32-bit field defines the number assigned to the first byte of data contained in this segment.

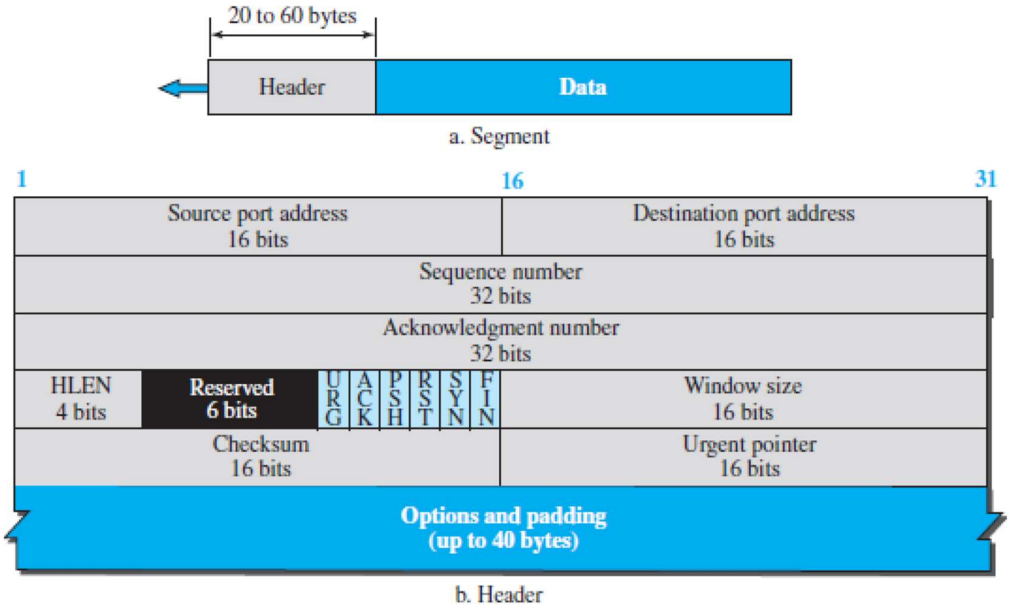

Fig4.2.1: TCP segment.

[Source : "Data Communications and Networking" by Behrouz A. Forouzan,Page-748]

TCP is a stream transport protocol.

To ensure connectivity, each byte to be transmitted is numbered. The sequence number tells the destination which byte in this sequence is the first byte in the segment. During connection establishment each party uses a random number generator to create an initial sequence number (ISN), which is usually different in each direction.

Acknowledgment number. This 32-bit field defines the byte number that the receiver of the segment is expecting to receive from the other party. If the receiver of the segment has successfully received byte number x from the other party, it returns x +1 as the acknowledgment number. Acknowledgment and data can be piggybacked together.

Header length. This 4-bit field indicates the number of 4-byte words in the TCP header. The length of the header can be between 20 and 60 bytes.

Control. This field defines 6 different control bits or flags.These bits enable flow control, connection establishment and termination, and the mode of data transfer in TCP.

Window size. This field defines the window size of the sending TCP in bytes. The length of this field is 16 bits, which means that the maximum size of the window is 65,535 bytes.

This value is referred to as the receiving window(rwnd) and is determined by the receiver. The sender must obey the dictation of the receiver in this case.

Checksum. This 16-bit field contains the checksum (error detection).The calculation of the checksum for TCP is important. The same pseudo header, serving the same purpose, is added to the segment.

Urgent pointer. This 16-bit field, which is valid only if the urgent flag is set, is used when the segment contains urgent data. It defines a value that must be added to the sequence number to obtain the number of the last urgent byte in the data section of the segment.

Options. There can be up to 40 bytes of optional information in the TCP header.

Encapsulation

A TCP segment encapsulates the data received from the application layer.