Syllabus
Transport Layer: Duties of Transport layer, Transport connection, OSI Transport protocol, TCP, UDP.

Transport Layer

***Transport Layer*:** Ensures that data is delivered error free, in sequence and with no loss, duplications or corruption. This layer also repackages data by assembling long messages into lots of smaller messages for sending, and repackaging the smaller messages into the original larger message at the receiving end.

The transport level provides end-to-end communication between processes executing on different machines. Although the services provided by a transport protocol are similar to those provided by a data link layer protocol, there are several important differences between the transport and lower layers:

1. User Oriented. Application programmers interact directly with the transport layer, and from the programmer's perspective, the transport layer is the ``network''. Thus, the transport layer should be oriented more towards user services than simply reflect what the underlying layers happen to provide. (Similar to the beautification principle in operating systems.)

2. Negotiation of Quality and Type of Services. The user and transport protocol may need to negotiate as to the quality or type of service to be provided. Examples? A user may want to negotiate such options as: throughput, delay, protection, priority, reliability, etc.

3. Guarantee Service. The transport layer may have to overcome service deficiencies of the lower layers (e.g. providing reliable service over an unreliable network layer).

4. Addressing becomes a significant issue. That is, now the user must deal with it; before it was buried in lower levels.

Two solutions:

- Use well-known addresses that rarely if ever change, allowing programs to ``wire in'' addresses. For what types of service does this work? While this works for services that are well established (e.g., mail, or telnet), it doesn't allow a user to easily experiment with new services.
- Use a name server. Servers register services with the name server, which clients contact to find the transport address of a given service.

5. Connection establishment. Transport level protocols go through three phases: establishing, using, and terminating a connection. For data gram-oriented protocols, opening a connection simply allocates and initializes data structures in the operating system kernel.

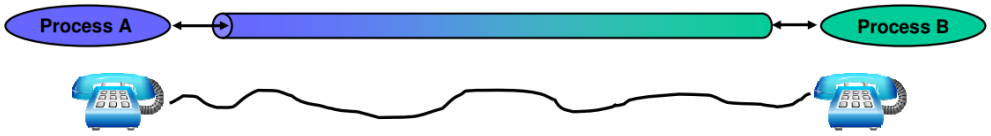
Connection oriented protocols often exchanges messages that negotiate options with the remote peer at the time a connection is opened. Establishing a connection may be tricky because of the possibility of old or duplicate packets.

Transport Layer Protocols:

A transport layer protocol is either **connection-oriented** or **connection-less**.

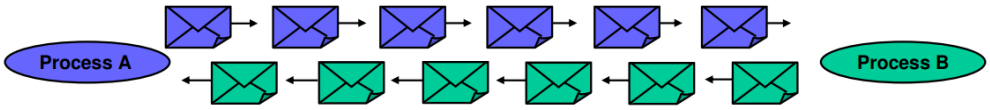
Connection-oriented transport protocols:

The peers establish a connection prior to a data exchange. This is similar to a telephone line that needs setting up a connection prior to a conversation.



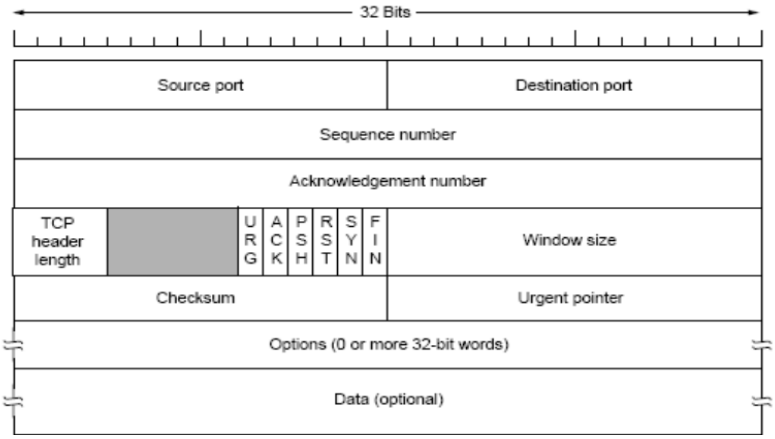
Connection-less transport protocols:

The peers send packets without a prior connection establishment. This is similar to the traditional postal service.



TCP and the Header Format

TCP (Transmission Control Protocol) was specifically designed to provide a reliable end-to-end byte stream over an unreliable internet work. An internet work differs from a single network because different parts may have wildly different topologies, bandwidths, delays, packet sizes, and other parameters. TCP was designed to dynamically adapt to properties of the internet work and to be robust in the face of many kinds of failures.



Header Format: -

The segment consists of a 20- to 60-byte header, 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.

1. 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. This serves the same purpose as the source port address in the UDP header.

2. 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. This serves the same purpose as the destination port address in the UDP header.

3. Sequence number. This 32-bit field defines the number assigned to the first byte of data contained in this segment. As we said before, 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 comprises 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.

4. 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 defines x + I as the acknowledgment number. Acknowledgment and data can be piggybacked together.

5. 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. Therefore, the value of this field can be between 5 (5 x 4 =20) and 15 (15 x 4 =60).

6. Reserved. This is a 6-bit field reserved for future use.

7. Control. This field defines 6 different control bits or flags.

8. Window size. This field defines the size of the window, in bytes, that the other party must maintain. Note that the length of this field is 16 bits, which means that the maximum size of the window is 65,535 bytes. This value is normally 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.

9. Checksum. This 16-bit field contains the checksum. The calculation of the checksum for TCP follows the same procedure as the one described for UDP. However, the inclusion of the checksum in the UDP datagram is optional, whereas the inclusion of the checksum for TCP is mandatory. The same pseudoheader, serving the same purpose, is added to the segment. For the TCP pseudoheader, the value for the protocol field is 6.

10. Urgent pointer. This l6-bit field, which is valid only if the urgent flag is set, is used when the segment contains urgent data. It defines the number that must be added to the sequence number to obtain the number of the last urgent byte in the data section of the segment.

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

TCP Connection

TCP is connection-oriented. A connection-oriented transport protocol establishes a virtual path between the source and destination. All the segments belonging to a message are then sent over this virtual path.

Using a single virtual pathway for the entire message facilitates the acknowledgment process as well as retransmission of damaged or lost frames. You may wonder how TCP, which uses the services of IP, a connectionless protocol, can be connection-oriented. The point is that a TCP connection is virtual, not physical. TCP operates at a higher level.

TCP uses the services of IP to deliver individual segments to the receiver, but it controls the connection itself. If a segment is lost or corrupted, it is retransmitted. Unlike TCP, IP is unaware of this retransmission. If a segment arrives out of order, TCP holds it until the missing segments arrive; IP is unaware of this reordering.

In TCP, connection-oriented transmission requires three phases: connection establishment, data transfer, and connection termination. TCP transmits data in full-duplex mode. When two TCPs in two machines are connected, they are able to send segments to each other simultaneously. This implies that each party must initialize communication and get approval from the other party before any data are transferred.

Three-Way Handshaking The connection establishment in TCP is called three-way handshaking.

Connection Establishment

The process starts with the server. The server program tells its TCP that it is ready to accept a connection. This is called a request for a *passive open.* Although the server TCP is ready to accept any connection from any machine in the world, it cannot make the connection itself.

The client program issues a request for an *active open.* A client that wishes to connect to an open server tells its TCP that it needs to be connected to that particular server.

The three steps in this phase are as follows.

1. The client sends the first segment, a SYN segment, in which only the SYN flag is set. This segment is for synchronization of sequence numbers. It consumes one sequence number. When the data transfer starts, the sequence number is incremented by 1. We can say that the SYN segment carries no real data.
2. The server sends the second segment, a SYN + ACK segment, with 2 flag bits set: SYN and ACK. This segment has a dual purpose. It is a SYN segment for communication in the other direction and serves as the acknowledgment for the SYN segment. It consumes one sequence number.
3. The client sends the third segment. This is just an ACK segment. It acknowledges the receipt of the second segment with the ACK flag and acknowledgment number field. Note that the sequence number in this segment is the same as the one in the SYN segment; the ACK segment does not consume any sequence numbers.

Data Transfer

After connection is established, bidirectional data transfer can take place. The client and server can both send data and acknowledgments. The data travels in the same direction as an acknowledgment are carried on the same segment. The acknowledgment is piggybacked with the data.

Connection Termination

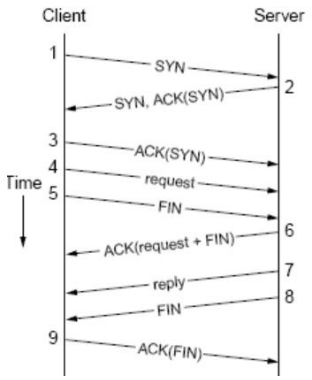
Any of the two parties involved in exchanging data (client or server) can close the connection, although it is usually initiated by the client. Most implementations today allow two options for connection termination: three-way handshaking and four-way handshaking with a half-close option.

1. In a normal situation, the client TCP, after receiving a close command from the client process, sends the first segment, a FIN segment in which the FIN flag is set. Note that a FIN segment can include the last chunk of data sent by the client, or it can be just a control segment.
2. The server TCP, after receiving the FIN segment, informs its process of the situation and sends the second segment, a FIN + ACK segment, to confirm the receipt of the FIN segment from the client and at the same time to announce the closing of the connection in the other direction. This segment can also contain the last chunk of data from the server. If it does not carry data, it consumes only one sequence number.
3. The client TCP sends the last segment, an ACK segment, to confirm the receipt of the FIN segment from the TCP server. This segment contains the acknowledgment number, which is 1 plus the sequence number received in the FIN segment from the server. This segment cannot carry data and consumes no sequence numbers.

RPC Using TPC

The normal sequence of packets for doing an RPC over TCP is shown in Fig. Nine packets are required in the best case. The nine packets are as follows:

1. The client sends a SYN packet to establish a connection.
2. The server sends an ACK packet to acknowledge the SYN packet.
3. The client completes the three-way handshake.
4. The client sends the actual request.
5. The client sends a FIN packet to indicate that it is done sending.
6. The server acknowledges the request and the FIN.
7. The server sends the reply back to the client.
8. The server sends a FIN packet to indicate that it is also done.
9. The client acknowledges the server's FIN.



UDP – User Datagram Protocol

UDP provides a connectionless service. This means that each user datagram sent by UDP is an independent datagram. There is no relationship between the different user datagrams even if they are coming from the same source process and going to the same destination program.

The user datagrams are not numbered. Also, there is no connection establishment and no connection termination, as is the case for TCP. This means that each user datagram can travel on a different path.

One of the ramifications of being connectionless is that the process that uses UDP cannot send a stream of data to UDP and expect UDP to chop them into different related user datagrams. Instead 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

UDP is a very simple, unreliable transport protocol. There is no flow control and hence no window mechanism. The receiver may overflow with incoming messages. There is no error control mechanism in UDP except for the checksum. This means that the sender does not know if a message has been lost or duplicated. When the receiver detects an error through the checksum, the user datagram is silently discarded. The lack of flow control and error control means that the process using UDP should provide these mechanisms.

Encapsulation and Decapsulation

To send a message from one process to another, the UDP protocol encapsulates and decapsulates messages in an IP datagram.

Queuing

We have talked about ports without discussing the actual implementation of them. In UDP, queues are associated with ports. At the client site, when a process starts, it requests a port number from the operating system. Some implementations create both an incoming and an outgoing queue associated with each process. Other implementations create only an incoming queue associated with each process.

Use of UDP

The following lists some uses of the UDP protocol:

1. UDP is suitable for a process that requires simple request-response communication with little concern for flow and error control.
2. UDP is suitable for a process with internal flow and error control mechanisms. For example, the Trivial File Transfer Protocol (TFTP) process includes flow and error control. It can easily use UDP.
3. UDP is a suitable transport protocol for multicasting. Multicasting capability is embedded in the UDP software but not in the TCP software.
4. UDP is used for management processes such as SNMP.
5. UDP is used for some route updating protocols such as Routing Information Protocol.