TCP/IP PROTOCOL SUITE

M. Lutfar Rahman

Introduction

In distributed processing and computer networking entities in different systems need to communicate. An entity is anything capable of sending and receiving information and a system is a physical object that contains one or more entities. User application programs, file transfer packages, database management systems, electronic mail facilities are the examples of entities, and terminals and computers are examples of physical systems.

Two entities must understand each other for successful communication between them. For information interchange between two entities, the entities must conform to some mutually acceptable set of conventions. This set of conventions is known as protocol. Protocol is a set of rules governing the exchange of data between two entities. Elements of a protocol include: data format, coding, signal levels, control information, error handling, speed of transmission and some other items necessary for meaningful exchange of information between two entities.

The TCP/IP protocol suite is an outgrowth of the development of ARPANET (American Research Project Agency Network). As time went on and as the ARPANET grew into the ARPA Internet, which included many subnets (such as MILNET, BITNET, CSENT etc.) LANs, several satellite channels and packet radio networks, the end to end reliability of the subnets declined. As a result a major change in the transport layer becomes necessary for unreliable subnets. This development led to the introduction of TCP (Transmission Control Protocol) which was designed for unreliable subnets. Associated with TCP a new network protocol named IP (Intent Protocol) was introduced. Currently TCP/IP is not only used in the ARPANET and ARPA Internet, but in many other commercial networks.

The TCP transport layer accepts long messages from the user processes and brakes them up into pieces not exceeding 64K bytes and sends a piece as a separate datagram. The network does not guarantee the proper delivery of a datagram and the TCP should retransmit the datagram if necessary. The datagrams may arrive at the receiver in the wrong order and the TCP assemble them in the proper sequence. Therefore, every datagram transmitted by the TCP should have its own sequence number.

Both the TCP/IP Protocol and the well known OSI (Open System Interconnection) protocol model of ISO (International Standards Organisation) deal with communication between many heterogeneous computers. Both have many similarities. However, there are some difference between the TCP/IP protocol suite and the ISO model.

The TCP/IP protocol suite gives equal importance on connectionless and connection oriented services. A connectionless service, such as datagram service, is one in which data are transferred from one entity to another without the prior establishment of a connection. On the other, hand in connection-oriented services data path or connection is set up before exchange of data between two entities. Presently the OSI model is used for connection oriented services. It is expected that future version of the OSI model will incorporate connectionless services.

TCP/IP PROTOCOL Architecture

The TCP/IP protocol architecture is based on communication that involves the agents processes, hosts and networks. Processes are the entities that communicate. A host or a station supports multiple processes that execute on hosts. The hosts are attached to the networks and the communication between processes takes place through networks. A network is concerned with routing data between hosts, when the hosts agree to process them.

The TCP/IP protocol suite is organised in four layers: network access layer, Internet layer, host-host layer and process/application layer (Fig. 1).

Process/Application Layer
Host-Host Layer
Internet Layer
Network Access Layer

Fig. 1 : TCP/IP Architecture

There are philosophical and practical difference between the OSI model and the TCP/IP Protocol suite.

The network access layer of TCP/IP protocol suite contains protocol for accessing communication network. This protocol between a communication node and a host (computer) attached to the node is called network access protocol. This protocol routes data between hosts attached to the same network. Flow control, error control between hosts and various service features, such as priority and security, may also be provided by this protocol. The network layer entity is typically invoked by an entity in the Internet or host-host layer, but it may also be invoked by the process/application layer directly.

The internet layer offers the basic functions required for the interconnection of dissimilar networks and hardware. It provides a datagram service for interconnecting a source computer and a destination computer through one or more dissimilar intermediate networks. The internet layer consists of procedures required to allow data to move through multiple networks between hosts. This protocol providing routing function is usually implemented in gateways or hosts and is known as the Internet Protocol or IP. A gateway, which is a processor, connects two networks and relays data between networks using an internet protocol.

The host-to-host layer permits the establishment of reliable connections between application programs resident in host computers. The protocols of the host-to-host layer are responsible to deliver data between two processes on different host computers. Other possible services of this level include error and flow control and the ability to deal with control signals not associated with a logical data connection. Four general types of protocol of host-to-host layer are: a reliable connection oriented data protocol, a datagram protocol, a speech protocol, and a real-time data protocol.

Many data processing applications use the reliable connection-oriented data protocol. This protocol is characterised by the requirement for reliable sequenced delivery of data. The datagram protocol with low overhead may be appropriate for applications that implement their own connection-oriented functionality. The speech protocol is characterised by the need for steady stream of data with minimum variation of delay. The real-time data protocol should posses the characteristics of both reliable connection-oriented protocol and the speech protocol.

The process/application layer contains protocols for sharing resources between computers and for accessing remote computers. File transfer protocol (FTP), simple mail transfer protocol (SMTP) and Telnet are some examples of the process/application layer protocols.

Operation of TCP and IP

A communication path between two computers may consist of multiple networks; the constituent networks are usually called subnetworks (Fig. 2).

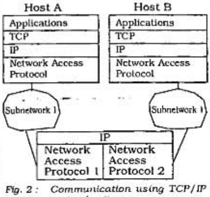

Fig. 2 : Communication using TCP/IP protocol suite

Some network access protocol is used to connect a computer to a subnetwork. This protocol enables a computer (host) to send data to another computer across the subnetworks. In the case of a host in another subnetwork, the data are sent through a router. The IP acts as a relay to move a block of date from one host to another through one or more routers. It is implemented in all the end systems and the routers. The TCP keeps track of the blocks of data so that they are delivered reliably to the appropriate applications. The TCP, thus, should be implemented on the end systems only.

Every entity in the communication networks must have a unique address. A host on a subnetwork must have global internet address so that the data can be delivered to the proper host. A process in a host has a unique address (also called port) within the host, and the host-to-host protocol can deliver data to the proper process. The TCP offers standardised ports to software resident in a host; for example, port 21 and port 23 are used by the FTP and the Telnet applications respectively.

How the data transfer operation is actually performed? To answer this question let us suppose that a process associated with port-1 at host A, wants to send a message to a process associated with port-2 at host B. The process at A hands data to the TCP with instructions to send it to port 2 at B. The TCP then hands the data to the IP with instruction to send it to B. The IP hands the data to the network access layer with instruction to send it to the router. For the operations described above, control information and user data are transmitted (Fig. 3).

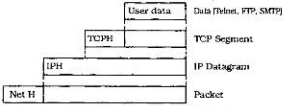

Fig. 3 : TCP/IP Protocol data Units

In the sending process the TCP may beak the data block into smaller pieces to make it more manageable. The TCP adds control information to each of these pieces forming a TCP segment. This control information is called TCP header (TCPH). The host B at the other end uses the header. This header should contain destination port address, sequence number and checksum. The destination port address identifies the process to whom the data should be delivered. The sequence number identifies the pieces of data sequentially so that if they arrive out of order, the TCP entity at B can reorder them properly. The checksum identifies error in transmission.

The TCP hands each segment over to the IP with instruction to send it to B. The IP appends a control header (IPH) to each segment forming an IP datagram. The IP header (IPH) should contain the address of the destination host B. Thus it is possible to transmit the segments across one or more subnetworks and routers.

Each IP datagram is presented to the network access layer for transmission across the first subnetwork in its journey to the destination. The network access layer appends its own header, NetH (called packet header), forming a frame or packet. The packet is then transmitted through the subnetwork. The packet header should contain destination subnet address and requests for some services that may include priority.

The router stripes off the packet header, examines the IP header and on the basis of destination address in the IP header, directs the datagram across subnetwork 2 for host B. In this process a new network access header is appended to the datagram.

The reverse process occurs when the packet is received by host B. At each layer of host B, the corresponding header is removed and the remainder is passed on to the next higher layer. This process continues until the original data are delivered to the destination process of host B.

Protocol Interfaces and Applications

A layer in the TCP/IP protocol suite interacts with its immediate adjacent layers At the source host, each layer directs data down to the next lower layer toward the network access layer and at the distinction, each layer delivers data to the next higher layer. However, the use of each layer is not required by the architecture. It is possible to develop applications that directly invoke the services of any one of the layers (Fig. 4).

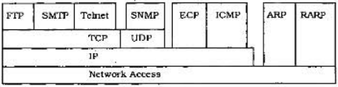

Fig. 4 : TCP/IP applications and interfaces

Some applications do not use the TCP layer. An application called SNMP (simple network management protocol) is a management and administration program for non-uniform networks. It uses an alternative user to user protocol called UDP (user datagram protocol), which provides a connectionless transfer mode between terminal users as an alternative to the TCP. Some applications (such as EGP - external gateway protocol, ICMP - Internet control message protocol) use IP directly. Some applications do not involve internetworking and thus do not require TCP at all; these applications (such as ARP - address resolution protocol, RARP - reverse address resolution protocol) directly invoke the network access layer. In fact a variety of other applications and processes make use of the TCP/IP architecture.

The common applications namely STMT, FTP and Telnet are discussed below. The SMTP provides a basic electronic mail facility. This application provides mechanism for transferring messages between different hosts. The SMTP includes mailing lists, return receipts and forwarding features. Some local editing or native electronic mail facility is required for creating messages for the SMTP. The SMTP accepts the message created by an appropriate facility and makes use of the TCP to send it to an SMTP module on another host. The receiving SMTP module stores the message in a user's mail box with the help of a local electronic mail package.

The FTP is used to send files (both text and binary) from one system to another. In response to the request for a file transfer, FTP sets up a TCP control connection to the target system and exchange control messages. The user ID, password and file specifications are transmitted through this connection. Once a file transfer is accepted, a second TCP connection is set up for transfer of data. The file is transferred directly through the data connection without any overhead of headers or control information. The control connection is then used to signal the completion of the transfer operation and to accept new file transfer commands.

The Telnet provides remote logon capability. It enables a user having a personal computer or a terminal to logon to a remote computer. The personal computer then functions as if it is directly connected to the remote computer. The traffic between the user and the remote computer is carried on a TCP connection. ❖