Introduction to TCP/IP

نام نویسنده مقاله: فرشته

SOROOSH450@YAHOO.COM

دانشگاه صنعتی شیراز

Summary: TCP and IP were developed by a Department of Defense (DOD) research project to connect a number different networks designed by different vendors into a network of networks (the "Internet"). It was initially successful because it delivered a few basic services that everyone needs (file transfer, electronic mail, remote logon) across a very large number of client and server systems. Several computers in a small department can use TCP/IP (along with other protocols) on a single LAN. The IP component provides routing from the department to the enterprise network, then to regional networks, and finally to the global Internet. On the battlefield a communications network will sustain damage, so the DOD designed TCP/IP to be robust and automatically recover from any node or phone line failure. This design allows the construction of very large networks with less central management. However, because of the automatic recovery, network problems can go undiagnosed and uncorrected for long periods of time.

- **IP** - is responsible for moving packet of data from node to node. IP forwards each packet based on a four byte destination address (the IP number). The Internet authorities assign ranges of numbers to different organizations. The organizations assign groups of their numbers to departments. IP operates on gateway machines that move data from department to organization to region and then around the world.
- **TCP** - is responsible for verifying the correct delivery of data from client to server. Data can be lost in the intermediate network. TCP adds support to detect errors or lost data and to trigger retransmission until the data is correctly and completely received.
- **Sockets -** is a name given to the package of subroutines that provide access to TCP/IP on most systems.

Background

The Internet protocols are the world's most popular open-system (nonproprietary) protocol suite because they can be used to communicate across any set of interconnected networks and are equally well suited for LAN and WAN communications. The Internet protocols consist of a suite of communication protocols, of which the two best known are the Transmission Control Protocol (TCP) and the Internet Protocol (IP). The Internet protocol suite not only includes lower-layer protocols (such as TCP and IP), but it also specifies common applications such as electronic mail, terminal emulation, and file transfer.

Network of Lowest Bidders

The Army puts out a bid on a computer and DEC wins the bid. The Air Force puts out a bid and IBM wins. The Navy bid is won by Unisys. Then the President decides to invade Grenada and the armed forces discover that their computers cannot talk to each other. The DOD must build a "network" out of systems each of which, by law, was delivered by the lowest bidder on a single contract. The Internet Protocol was developed to create a Network of Networks (the "Internet"). Individual machines are first connected to a LAN (Ethernet or Token Ring). TCP/IP shares the LAN with other uses (a Novell file server, Windows for Workgroups peer systems). One device provides the TCP/IP connection between the LAN and the rest of the world. To insure that all types of systems from all vendors can communicate, TCP/IP is absolutely standardized on the LAN. However, larger networks based on long distances and phone lines are more volatile. The original design of TCP/IP as a Network of Networks fits nicely within the current technological uncertainty. TCP/IP data can be sent across a LAN, or it can be carried within an internal corporate SNA network, or it can piggyback on the cable TV service. Furthermore, machines connected to any of these networks can communicate to any other network through gateways supplied by the network vendor.

Addresses

Each technology has its own convention for transmitting messages between two machines within the same network. On a LAN, messages are sent between machines by supplying the six byte unique identifier (the "MAC" address). In an SNA network, every machine has Logical Units with their own network address. DECNET, Appletalk, and Novell IPX all have a scheme for assigning numbers to each local network and to each workstation attached to the network. On top of these local or vendor specific network addresses, TCP/IP assigns a unique number to every workstation in the world. This "IP number" is a four byte value that, by convention, is expressed by converting each byte into a decimal number (0 to 255) and separating the bytes with a period. For example, the PC Lube and Tune server is 130.132.59.234. It is still possible for almost anyone to get assignment of a number for a small "Class C" network in which the first three bytes identify the network and the last byte identifies the individual computer. The author followed this procedure and was assigned the numbers 192.35.91.* for a network of computers at his house. Larger organizations can get a "Class B" network where the first two bytes identify the network and the last two bytes identify each of up to 64 thousand individual workstations. Yale's Class B network is 130.132, so all computers with IP address 130.132.*.* are connected through Yale.

The organization then connects to the Internet through one of a dozen regional or specialized network suppliers. The network vendor is given the subscriber network number and adds it to the routing configuration in its own machines and those of the other major network suppliers. There is no mathematical formula that translates the numbers 192.35.91 or 130.132 into "Yale University" or "New Haven, CT." The machines that manage large regional networks or the central Internet routers managed by the National Science Foundation can only locate these networks by looking each network number up in a table. There are potentially thousands of Class B networks, and millions of Class C networks, but computer memory costs are low, so the tables are reasonable. Customers that connect to the Internet, even customers as large as IBM, do not need to maintain any information on other networks. They send all external data to the regional carrier to which they subscribe, and the regional carrier maintains the tables and does the appropriate routing.

Subnets

Although the individual subscribers do not need to tabulate network numbers or provide explicit routing, it is convenient for most Class B networks to be internally managed as a much smaller and simpler version of the larger network organizations. It is common to subdivide the two bytes available for internal assignment into a one byte department number and a one byte workstation ID. The enterprise network is built using commercially available TCP/IP router boxes. Each router has small tables with 255 entries to translate the one byte department number into selection of a destination Ethernet connected to one of the routers.

A Uncertain Path

Every time a message arrives at an IP router, it makes an individual decision about where to send it next. There is concept of a session with a preselected path for all traffic. Consider a company with facilities in New York, Los Angeles, Chicago and Atlanta. It could build a network from four phone lines forming a loop (NY to Chicago to LA to Atlanta to NY). A message arriving at the NY router could go to LA via either Chicago or Atlanta. The reply could come back the other way.

How does the router make a decision between routes? There is no correct answer. Traffic could be routed by the "clockwise" algorithm (go NY to Atlanta, LA to Chicago). The routers could alternate, sending one message to Atlanta and the next to Chicago. More sophisticated routing measures traffic patterns and sends data through the least busy link. If one phone line in this network breaks down, traffic can still reach its destination through a roundabout path. After losing the NY to Chicago line, data can be sent NY to Atlanta to LA to Chicago. This provides continued service though with degraded performance. This kind of recovery is the primary design feature of IP. The loss of the line is immediately detected by the routers in NY and Chicago, but somehow this information must be sent to the other nodes. Otherwise, LA could continue to send NY messages through Chicago, where they arrive at a "dead end." Each network adopts some Router Protocol which periodically updates the routing tables throughout the network with information about changes in route status. If the size of the network grows, then the complexity of the routing updates will increase as will the cost of transmitting them. Building a single network that covers the entire US would be unreasonably complicated. Fortunately, the Internet is designed as a Network of Networks. This means that loops and redundancy are built into each regional carrier. The regional network handles its own problems and reroutes messages internally. Its Router Protocol updates the tables in its own routers, but no routing updates need to propagate from a regional carrier to the NSF spine or to the other regions (unless, of course, a subscriber switches permanently from one region to another).

Undiagnosed Problems

IBM designs its SNA networks to be centrally managed. If any error occurs, it is reported to the network authorities. By design, any error is a problem that should be corrected or repaired. IP networks, however, were designed to be robust. In battlefield conditions, the loss of a node or line is a normal circumstance. Casualties can be sorted out later on, but the network must stay up. So IP networks are robust. They automatically (and silently) reconfigure themselves when something goes wrong. If there is enough redundancy built into the system, then communication is maintained.

In 1975 when SNA was designed, such redundancy would be prohibitively expensive, or it might have been argued that only the Defense Department could afford it. Today, however, simple routers cost no more than a PC. However, the TCP/IP design that, "Errors are normal and can be largely ignored," produces problems of its own. Data traffic is frequently organized around "hubs," much like airline traffic. One could imagine an IP router in Atlanta routing messages for smaller cities throughout the Southeast. The problem is that data arrives without a reservation. Airline companies experience the problem around major events, like the Super Bowl. Just before the game, everyone wants to fly into the city. After the game, everyone wants to fly out. Imbalance occurs on the network when something new gets advertised. Adam Curry announced the server at "mtv.com" and his regional carrier was swamped with traffic the next day. The problem is that messages come in from the entire world over high speed lines, but they go out to mtv.com over what was then a slow speed phone line.

Occasionally a snow storm cancels flights and airports fill up with stranded passengers. Many go off to hotels in town. When data arrives at a congested router, there is no place to send the overflow. Excess packets are simply discarded. It becomes the responsibility of the sender to retry the data a few seconds later and to persist until it finally gets through. This recovery is provided by the TCP component of the Internet protocol. TCP was designed to recover from node or line failures where the network propagates routing table changes to all router nodes. Since the update takes some time, TCP is slow to initiate recovery. The TCP algorithms are not tuned to optimally handle packet loss due to traffic congestion. Instead, the traditional Internet response to traffic problems has been to increase the speed of lines and equipment in order to say ahead of growth in demand. TCP treats the data as a stream of bytes. It logically assigns a sequence number to each byte. The TCP packet has a header that says, in effect, "This packet starts with byte 379642 and contains 200 bytes of data." The receiver can detect missing or incorrectly sequenced packets. TCP acknowledges data that has been received and retransmits data that has been lost. The TCP design means that error recovery is done end-to-end between the Client and Server machine. There is no formal standard for tracking problems in the middle of the network, though each network has adopted some ad hoc tools.

Need to Know

There are three levels of TCP/IP knowledge. Those who administer a regional or national network must design a system of long distance phone lines, dedicated routing devices, and very large configuration files. They must know the IP numbers and physical locations of thousands of subscriber networks. They must also have a formal network monitor strategy to detect problems and respond quickly. Each large company or university that subscribes to the Internet must have an intermediate level of network organization and expertise. A half dozen routers might be configured to connect several dozen departmental LANs in several buildings. All traffic outside the organization would typically be routed to a single connection to a regional network provider. However, the end user can install TCP/IP on a personal computer without any knowledge of either the corporate or regional network. Three pieces of information are required:

1. The IP address assigned to this personal computer
2. The part of the IP address (the subnet mask) that distinguishes other machines on the same LAN (messages can be sent to them directly) from machines in other departments or elsewhere in the world (which are sent to a router machine)
3. The IP address of the router machine that connects this LAN to the rest of the world.

In the case of the PCLT server, the IP address is 130.132.59.234. Since the first three bytes designate this department, a "subnet mask" is defined as 255.255.255.0 (255 is the largest byte value and represents the number with all bits turned on). It is a Yale convention (which we recommend to everyone) that the router for each department have station number 1 within the department network. Thus the PCLT router is 130.132.59.1. Thus the PCLT server is configured with the values:

- My IP address: 130.132.59.234
- Subnet mask: 255.255.255.0
- Default router: 130.132.59.1

The subnet mask tells the server that any other machine with an IP address beginning 130.132.59.* is on the same department LAN, so messages are sent to it directly. Any IP address beginning with a different value is accessed indirectly by sending the message through the router at 130.132.59.1 (which is on the departmental LAN).

Internet Protocol (IP)

The Internet Protocol (IP) is a network-layer (Layer 3) protocol that contains addressing information and some control information that enables packets to be routed. IP is documented in RFC 791 and is the primary network-layer protocol in the Internet protocol suite. Along with the Transmission Control Protocol (TCP), IP represents the heart of the Internet protocols. IP has two primary responsibilities: providing connectionless, best-effort delivery of datagrams through an internetwork; and providing fragmentation and reassembly of datagrams to support data links with different maximum-transmission unit (MTU) sizes.

IP Packet Format

The following discussion describes the IP packet fields :

- *Version*—Indicates the version of IP currently used.
- *IP Header Length (*IHL)—Indicates the datagram header length in 32-bit words.
- *Type-of-Service*—Specifies how an upper-layer protocol would like a current datagram to be handled, and assigns datagrams various levels of importance.
- *Total Length*—Specifies the length, in bytes, of the entire IP packet, including the data and header.
- *Identification*—Contains an integer that identifies the current datagram. This field is used to help piece together datagram fragments.
- *Flags*—Consists of a 3-bit field of which the two low-order (least-significant) bits control fragmentation. The low-order bit specifies whether the packet can be fragmented. The middle bit specifies whether the packet is the last fragment in a series of fragmented packets. The third or high-order bit is not used.
- *Fragment Offset*—Indicates the position of the fragment's data relative to the beginning of the data in the original datagram, which allows the destination IP process to properly reconstruct the original datagram.
- *Time-to-Live*—Maintains a counter that gradually decrements down to zero, at which point the datagram is discarded. This keeps packets from looping endlessly.
- *Protocol*—Indicates which upper-layer protocol receives incoming packets after IP processing is complete.
- *Header Checksum*—Helps ensure IP header integrity.
- S*ource Address*—Specifies the sending node.
- *Destination Address*—Specifies the receiving node.
- *Options*—Allows IP to support various options, such as security.
- *Data*—Contains upper-layer information.

IP Addressing

As with any other network-layer protocol, the IP addressing scheme is integral to the process of routing IP datagrams through an internetwork. Each IP address has specific components and follows a basic format. These IP addresses can be subdivided and used to create addresses for subnetworks.Each host on a TCP/IP network is assigned a unique 32-bit logical address that is divided into two main parts: the network number and the host number. The network number identifies a network and must be assigned by the Internet Network Information Center (InterNIC) if the network is to be part of the Internet. An Internet Service Provider (ISP) can obtain blocks of network addresses from the InterNIC and can itself assign address space as necessary. The host number identifies a host on a network and is assigned by the local network administrator.

IP Address Format

The 32-bit IP address is grouped eight bits at a time, separated by dots, and represented in decimal format (known as *dotted decimal notation*). Each bit in the octet has a binary weight (128, 64, 32, 16, 8, 4, 2, 1). The minimum value for an octet is 0, and the maximum value for an octet is 255.

IP Address Classes

IP addressing supports five different address classes: A, B,C, D, and E. Only classes A, B, and C are available for commercial use. The left-most (high-order) bits indicate the network class. The class of address can be determined easily by examining the first octet of the address. In an IP address of 172.31.1.2, for example, the first octet is 172. Because 172 falls between 128 and 191, 172.31.1.2 is a Class B address.

IP Subnet Addressing

IP networks can be divided into smaller networks called subnetworks (or subnets). Subnetting provides the network administrator with several benefits, including extra flexibility, more efficient use of network addresses, and the capability to contain broadcast traffic (a broadcast will not cross a router).Subnets are under local administration. As such, the outside world sees an organization as a single network and has no detailed knowledge of the organization's internal structure.A given network address can be broken up into many subnetworks. For example, 172.16.1.0, 172.16.2.0, 172.16.3.0, and 172.16.4.0 are all subnets within network 171.16.0.0. (All 0s in the host portion of an address specifies the entire network.)

IP Subnet Mask

A subnet address is created by "borrowing" bits from the host field and designating them as the subnet field. The number of borrowed bits varies and is specified by the subnet mask. Subnet masks use the same format and representation technique as IP addresses. The subnet mask, however, has binary 1s in all bits specifying the network and subnetwork fields, and binary 0s in all bits specifying the host field. Various types of subnet masks exist for Class B and C subnets.The default subnet mask for a Class B address that has no subnetting is 255.255.0.0, while the subnet mask for a Class B address 171.16.0.0 that specifies eight bits of subnetting is 255.255.255.0. The reason for this is that eight bits of subnetting or $2^8 - 2$ (1 for the network address and 1 for the broadcast address) = 254 subnets possible, with $2^8 - 2 = 254$ hosts per subnet.The subnet mask for a Class C address 192.168.2.0 that specifies five bits of subnetting is 255.255.255.248.With five bits available for subnetting, $2^5 - 2 = 30$ subnets possible, with
$2^3 - 2 = 6$ hosts per subnet.

Transmission Control Protocol (TCP)

The TCP provides reliable transmission of data in an IP environment. TCP corresponds to the transport layer (Layer 4) of the OSI reference model. Among the services TCP provides are stream data transfer, reliability, efficient flow control, full-duplex operation, and multiplexing.With stream data transfer**,** TCP delivers an unstructured stream of bytes identified by sequence numbers. This service benefits applications because they do not have to chop data into blocks before handing it off to TCP. Instead, TCP groups bytes into segments and passes them to IP for delivery. TCP offers reliability by providing connection-oriented, end-to-end reliable packet delivery through an internetwork. It does this by sequencing bytes with a forwarding acknowledgment number that indicates to the destination the next byte the source expects to receive. Bytes not acknowledged within a specified time period are retransmitted. The reliability mechanism of TCP allows devices to deal with lost, delayed, duplicate, or misread packets. A time-out mechanism allows devices to detect lost packets and request retransmission.TCP offers efficient flow control, which means that, when sending acknowledgments back to the source, the receiving TCP process indicates the highest sequence number it can receive without overflowing its internal buffers.Full-duplex operation means that TCP processes can both send and receive at the same time.Finally, TCP's multiplexing means that numerous simultaneous upper-layer conversations can be multiplexed over a single connection.

TCP Connection Establishment

To use reliable transport services, TCP hosts must establish a connection-oriented session with one another. Connection establishment is performed by using a "three-way handshake" mechanism. A three-way handshake synchronizes both ends of a connection by allowing both sides to agree upon initial sequence numbers. This mechanism also guarantees that both sides are ready to transmit data and know that the other side is ready to transmit as well. This is necessary so that packets are not transmitted or retransmitted during session establishment or after session termination.Each host randomly chooses a sequence number used to track bytes within the stream it is sending and receiving. Then, the three-way handshake proceeds in the following manner:The first host (Host A) initiates a connection by sending a packet with the initial sequence number (X) and SYN bit set to indicate a connection request. The second host (Host B) receives the SYN, records the sequence number X, and replies by acknowledging the SYN (with an ACK = X + 1). Host B includes its own initial sequence number (SEQ = Y). An ACK = 20 means the host has received bytes 0 through 19 and expects byte 20 next. This technique is called *forward acknowledgment*. Host A then acknowledges all bytes Host B sent with a forward acknowledgment indicating the next byte Host A expects to receive (ACK = Y + 1). Data transfer then can begin.

Positive Acknowledgment and Retransmission (PAR)

A simple transport protocol might implement a reliability-and-flow-control technique where the source sends one packet, starts a timer, and waits for an acknowledgment before sending a new packet. If the acknowledgment is not received before the timer expires, the source retransmits the packet. Such a technique is called *positive acknowledgment and retransmission* (PAR). By assigning each packet a sequence number, PAR enables hosts to track lost or duplicate packets caused by network delays that result in premature retransmission. The sequence numbers are sent back in the acknowledgments so that the acknowledgments can be tracked.PAR is an inefficient use of bandwidth, however, because a host must wait for an acknowledgment before sending a new packet, and only one packet can be sent at a time.

TCP Sliding Window

A *TCP sliding window* provides more efficient use of network bandwidth than PAR because it enables hosts to send multiple bytes or packets before waiting for an acknowledgment. In TCP, the receiver specifies the current window size in every packet. Because TCP provides a byte-stream connection, window sizes are expressed in bytes. This means that a window is the number of data bytes that the sender is allowed to send before waiting for an acknowledgment. Initial window sizes are indicated at connection setup, but might vary throughout the data transfer to provide flow control. A window size of zero, for instance, means "Send no data." In a TCP sliding-window operation, for example, the sender might have a sequence of bytes to send (numbered 1 to 10) to a receiver who has a window size of five. The sender then would place a window around the first five bytes and transmit them together. It would then wait for an acknowledgment.The receiver would respond with an ACK = 6, indicating that it has received bytes 1 to 5 and is expecting byte 6 next. In the same packet, the receiver would indicate that its window size is 5. The sender then would move the sliding window five bytes to the right and transmit bytes 6 to 10. The receiver would respond with an ACK = 11, indicating that it is expecting sequenced byte 11 next. In this packet, the receiver might indicate that its window size is 0 (because, for example, its internal buffers are full). At this point, the sender cannot send any more bytes until the receiver sends another packet with a window size greater than 0.

TCP Packet Field Descriptions

The following descriptions summarize the TCP packet fields:

- *Source Port* and *Destination Port*—Identifies points at which upper-layer source and destination processes receive TCP services.
- *Sequence Number*—Usually specifies the number assigned to the first byte of data in the current message. In the connection-establishment phase, this field also can be used to identify an initial sequence number to be used in an upcoming transmission.
- *Acknowledgment Number*—Contains the sequence number of the next byte of data the sender of the packet expects to receive.
- *Data Offset*—Indicates the number of 32-bit words in the TCP header.
- *Reserved*—Remains reserved for future use.

- *Flags*—Carries a variety of control information, including the SYN and ACK bits used for connection establishment, and the FIN bit used for connection termination.
- *Window*—Specifies the size of the sender's receive window (that is, the buffer space available for incoming data).
- *Checksum*—Indicates whether the header was damaged in transit.
- *Urgent Pointer*—Points to the first urgent data byte in the packet.
- *Options*—Specifies various TCP options.
- *Data*—Contains upper-layer information.

User Datagram Protocol (UDP)

The User Datagram Protocol (UDP) is a connectionless transport-layer protocol (Layer 4) that belongs to the Internet protocol family. UDP is basically an interface between IP and upper-layer processes. UDP protocol ports distinguish multiple applications running on a single device from one another. Unlike the TCP, UDP adds no reliability, flow-control, or error-recovery functions to IP. Because of UDP's simplicity, UDP headers contain fewer bytes and consume less network overhead than TCP.UDP is useful in situations where the reliability mechanisms of TCP are not necessary, such as in cases where a higher-layer protocol might provide error and flow control.UDP is the transport protocol for several well-known application-layer protocols, including Network File System (NFS), Simple Network Management Protocol (SNMP), Domain Name System (DNS), and Trivial File Transfer Protocol (TFTP).Source and destination ports contain the 16-bit UDP protocol port numbers used to demultiplex datagrams for receiving application-layer processes. A length field specifies the length of the UDP header and data. Checksum provides an (optional) integrity check on the UDP header and data.

Internet Protocols Application-Layer Protocols

The Internet protocol suite includes many application-layer protocols that represent a wide variety of applications, including the following:

- *File Transfer Protocol (FTP)*—Moves files between devices
- *Simple Network-Management Protocol (SNMP)*—Primarily reports anomalous network conditions and sets network threshold values
- *Telnet*—Serves as a terminal emulation protocol
- *X Windows*—Serves as a distributed windowing and graphics system used for communication between X terminals and UNIX workstations
- *Network File System (NFS), External Data Representation (XDR), and Remote Procedure Call (RPC)*—Work together to enable transparent access to remote network resources
- *Simple Mail Transfer Protocol (SMTP)*—Provides electronic mail services
- *Domain Name System (DNS)*—Translates the names of network nodes into network addresses

REFRENCES:

enternet ,www.ni.com,protocols internet
Mohsen SHekibafar,"comunication by I/O ",Transistor ,(of shiraz sanati univercity)5,82
مهندس فربد قابوسي, راهنماي جامع *LabVIEW*, Tehran,nashr afarang, page:528,82