Computer Networks

Computer Networks and the Internet

The network is a system of links that interconnect different sets of nodes that communicate with each other by information.
The internet is a network of networks. It means that the internet is basically an umbrella that merges all the networks together.
In computer networks, an end system is a device that is connected directly to the internet.
End systems = Hosts.
Packet switches forward packets of data between each other and between hosts or end devices.

Two types of packet switches:
- Routers
- Switches

In the same manner that a switch connects multiple devices to create a network, a router connects multiple switches, and their respective networks, to form an even larger network.
Communication links interconnect routers, switches, hosts, and end systems.

Protocols control the sending and receiving of data. Furthermore, a protocol defines the format and the order of messages exchanged between two or more communication entities. For the reason that protocols describe a standard way of doing things, there needs to be a body that defines those standards. For the internet, the internet standards, and protocols, that defining body is known as the Internet Engineering Task Force (IETF) and their standards are known as Request for Comments (RFC).
From the server's perspective, the internet is all about the delivery of information from one point in the network to another.
Types of access networks:
- Residential access networks
- Institutional access networks (schools, companies)
- Mobile access networks (Wifi, 4G/5G)

Transmission rate or access rate refers to the speed at which data can be transmitted over that network. It represents the amount of data that can be sent or received in a given amount of time.

Cable-Based Access

For the cable-based access, there is a special type of cable that goes through our homes and that type of cable contains data, and by data we mean anything.

Digital Subscriber Line

DSL is the connection to the internet through the telephone network. The digital subscriber line is widely used but it lacks high speed.

Home Networks

For the home network, we have one big router that all the internet connection is coming to our homes from, and by our devices we are connected to it through Wifi.

Wireless Access Networks

There are two types of wireless access networks:

- Wireless local area networks (WLANs)
- Wide area cellular access networks

For both types of wireless access networks, there's a base station or an access point to which the end devices transmit and receive data from.

Enterprise Networks

Enterprise networks are a mix of wired, and wireless technologies, connecting a mix of switches and routers to handle devices that would be connected to that enterprise network.

Data Center Networks

Data center networks connect a massive number of servers to each other and to the internet. Data centers contain a huge number of CPUs and RAMs and other components in order to function like a big supercomputer.

Network Core

How do we send packages through the Internet?
Internet nowadays works with the packet switching technique. The packet switching technique means that the host is sending a request or message through the Internet and the host is breaking that message into chunks of data, or in terms of computer networks, packets.

Two key network-core functions:

Forwarding

Forwarding is simply defined as the action applied by each router when a packet arrives at one of its interfaces. Inside each router, there is a routing table that stores the header values in binary numbers, and the output links that the packets will be forwarded. For example, if a message ends up with binary code 0111, then you need to check the table and see that for binary code 0111, the message needs to be forwarded to link number 2. Moreover, forwarding is a local connection that happens inside the router.

Routing

Routing is the process of path selection in any network. A computer network is made of nodes and links that connect those nodes. Communication between two nodes in an interconnected network can take place through many different paths. Routing is the process of selecting the best path using some predetermined rules. Moreover, routing is a global connection that happens on the top of the router.

Queueing in routing occurs when work arrives faster than it can be serviced.

From whom we are getting the Internet?
The Internet Service Providers (ISPs) provide us with the Internet.

How are ISPs connected with each other?
For the ISPs to connect with each other, we need some special routers called Internet Exchange Points (IXP), a physical location through which sets of ISPs connect with each other.

Application Layer

Packet loss is defined by having two hosts on one side and the destination on the other side. While the two hosts send packets, the packets will first go to the router, and since routers operate with a buffer inside of them, that buffer has a limited capacity, therefore if the buffer size exceeds, it means that we have a packet loss.

The latency formula explained in detail:

dproc (processing) - checks the bit errors.

dqeueue (queueing) - the time that the packets wait in the routers to be sent to the destination.

dproc and dequeue are always close to 0, therefore, in all the exercises where we calculate the latency, they are always taken as 0.

dtrans (transmission) - the time or the latency to deliver one packet from one side to another and it's generally calculated with the L/R formula where L is the file size and R is the capacity of the link.

dprop (propagation) - the amount of time it takes for the head of the signal to travel from one side to another and it's generally calculated with the d/s formula where d is the length of the critical link (generally in km) and s is the speed of light.

By adding all, processing delays, queueing delays, transmission, and propagation delays, we find the delay of the nodes.

If an Internet Service Provider offers us Internet with a bandwidth of 300Mbps that is the transmission rate of the link from my PC to the Internet, and when we measure the Internet connection to our homes, the bandwidth is usually in a lower value. That's because bandwidth is the theoretical maximum of the speed and when we measure the internet connection in our homes, that is the throughput, which is the momentum or the actual value. Both bandwidth and throughput are measured in bits per second.

Two types of applications:

Both client-server and peer-to-peer networks connect computers so that they can share resources from one computer to another, however, they differ from one another.

Client-Server - The server responds to the services that are requested by the client.

Peer-2-Peer - Each and every node can both request and respond to the services.

Computer networks are complex therefore everything is organized in layers. We divide the computer networks into 5 layers: Application layer, Transport layer, Network layer, Link layer, and Physical layer.

All hosts implement the 5 layers above. Routers, however, implement just 3 layers.

The application layer deals with processes. A process is a program running within a host. For example, when you send a message from Facebook on your PC, there is a Facebook process that sends this message to the other process somewhere else. Therefore, we have client processes that initiate communication, and given in our example, they send the message and also we have server processes that wait to be contacted, and given in our example, they wait to receive the message.

This whole process is in the client-server network. In peer-2-peer, each client and each server can have client and server requests at the same time.

How do the processes work so that hosts communicate with each other?
A process on one side communicates with a process from the other side by the process's socket number or port number. The socket number serves as a door for two processes to communicate with each other. The socket number is the combination of the address of the server and the port number. Socket number helps to recognize the address of the application to which data is to be sent using the address of the server and the port number.

What transport service does an application need?

Data integrity - means that some applications require 100% reliable data transfer. For example, Gmail, Data servers, File transfer, etc. Other applications such as Audio applications can tolerate some loss.

Timing - means that some applications require a low delay to be “effective”. For example, Internet telephony, interactive games, etc.

Throughput - means that some applications require a minimum amount of throughput to be “effective”. For example, multimedia.

Security - means that some applications require encryption and data integrity.

Services, Layering and Encapsulation
In the application layer, the message is first received and then sent through the network. In the transport layer, when the message comes from the above layer or the top layer, it is not a message anymore, it becomes a segment. From the transport layer, the segment goes to the network layer or the router, and the network layer does not understand the segment, therefore, it converts it to a datagram. Further, the datagram goes to the link layer which becomes a frame. Lastly, the frame goes to the physical layer, turns into bits, 1s, and 0s, and the message will go from one side to the other side.
When the messages are put inside each other, that process is called encapsulation. When the headers are removed, that process is called decapsulation.

A web page consists of a set of objects and the objects can be HTML files, JPEG images, audio files, etc.
The website consists of the hostname (name of the website) and the path name.
HTTP (Hypertext Transfer Protocol) is an application layer protocol that has two sides, the client side which makes the requests (the client from where we make the requests is the browser), and the server side which the web server sends objects in response to requests.
HTTP uses TCP (Transmission Control Protocol) which is a reliable protocol that transfers messages from one side to the other side. If we use TCP, we are 100% sure that the message will transfer from one side to another, however, if we use UDP, we cannot say that for sure.
HTTP is stateless meaning that HTTP does not store information about past clients. The requests from the clients are not stored.

There are two types of HTTP connections:

Non-persistent HTTP - means that when I'm sending a request to the other side, the connection is open, I send one object to the other side and the connection is closed.

Persistent HTTP - means that the server will open the connection for me to send multiple objects, thus, I send multiple requests, I download multiple files, and, in the end, the server will close the connection.

In non-persistent, we are sending one object whereas in persistent we are sending multiple objects.

RTT (Round Trip Time) is the time for a small packet to travel from client to server and back. If Round Trip Time is small in value, we get a fast response from the website whereas if it goes to hundreds of milliseconds, we need to wait.

The total time to get a response from the website for one single object:

Non-persistent HTTP response time = 2RTT+ file transmission time.

Overall, non-persistent HTTP requires 2RTTs.

GET method is used for sending data to the server and it include user data in the URL field of HTTP.

POST method defines user input sent from client to server in the entity body of HTTP.

HEAD method requests headers that would be returned if specified URL was requested with an HTTP GET method.

PUT method completely replaces file that exists at specified URL with content in entity body of HTTP POST request message.

HTTP response status codes:

200 OK - means that the request succeeded.

301 Moved Permanently - means that the object we requested moved somewhere else.

400 Bad Request - means a bad request.

404 Not Found - means that the requested document is not found on the server.

505 HTTP Version Not Supported.

Cookies store information which can be time to time dangerous. Mainly, websites use cookies to register the user's data in their database. For example, if I visit a specific website, when my HTTP request arrives at that specific website, the website's server creates an ID for me. By the ID that the website's server creates for me, they manage to store detail information about me in their database. All these type of information about myself are all cookies.

Cookies can be used for authorization, shopping carts, recommendations, user session states etc.

SMTP (Simple Mail Transfer Protocol) is the protocol for exchanging email messages.

There are three major components when it comes to emails:

User agents - the ones who send and receive email messages (Gmail, Yahoo, etc.)

Mail servers - the ones who receive the servers.

The SMTP protocol.

The detailed process of SMTP:
I'm sending a message, the message will go to my mail server and each of the different email addresses that are in the mail server, the SMTP protocol uses to exchange email messages between mail servers.

Even in the SMTP protocol, there is a client, which is sending the email and there is a server, which is receiving the email. So, the SMTP mail server can serve as a client and as a server as well.
TCP connection enables application programs and computing devices to exchange messages over a network. HTTP and SMTP both use TCP connection because it is reliable.

Differences between HTTP and SMTP?
In terms of how they work, HTTP is client pull, for example, if we go to a website, we pull the content from the website, whereas SMTP is client push, meaning that SMTP is forwarding and sending content.

When we want to download emails, we need to select one of the two protocols that are IMAP or POP protocol.
IMAP allows you to access your email wherever you are, from any device. When you read an email message using IMAP, you aren't downloading or storing it on your computer; instead, you're reading it from the email service. As a result, you can check your email from different devices, anywhere in the world: your phone, a computer, or a friend's computer.
POP works by contacting your email service and downloading all your new messages from it. Once they are downloaded onto your PC, they are deleted from the email service. This means that after the email is downloaded, it can only be accessed using the same computer. If you try to access your email from a different device, the messages that have been previously downloaded won't be available to you.

What do we use to identify hosts on the internet?
IP addresses are used to identify hosts on the internet but sometimes we even use host names or domain names. In fact, DNS (Domain Name System) turns domain names into IP addresses, which browsers use to load internet page addresses.

DNS servers are distributed. Why are not DNS servers centralized, all in one place?
It is very dangerous to have all DNS servers centralized or in one place together due to hackers and other security reasons. Another reason is the demand of people trying to access one DNS server that is in one place.

The way to provide multimedia content to users is by streaming. The term streaming refers to the continual transmission of audio and video files from a server to a client.
CDN (Content Distribution Network) is a geographically distributed and interconnected servers. They provide cached internet content from a network location closest to a user to speed up its delivery.

Transport Layer

In the transport layer, the communication is done between two processes in different hosts.
Two of the most important transport layer protocols are TCP (Transmission Control Protocol) and UDP (User Datagram Protocol).
TCP protocol is a reliable protocol and when we use it, we are 100% sure that the message that we send will arrive whereas UDP protocol is not 100% reliable. However, UDP is a faster protocol, and the three-way handshake method is not required in UDP.
The transport layer provides logical communication. It means that it provides two different processes to two different hosts that communicate with each other.

The difference between the Network and the Transport layer.
The main difference is that in the network layer, the communication is done between the hosts, and in the transport layer, the communication is done between the processes.

In the transport layer, processes are important because, on my PC, I don't have a single process, I have multiple processes. The transport layer has to do with process-to-process connection.

The differences between the TCP and the UDP protocol.
TCP is a reliable protocol meaning that if we send a message, we are 100% sure that the message will arrive or at least the destination will receive a notification that the message hasn't arrived. TCP has a conjunction control. Conjunction control means that the TCP protocol knows the network. TCP is connected-oriented.
UDP is unreliable meaning that we can have lots of packet losses. Regardless, UDP is very fast. In UDP there are no delay guarantees. UDP is connectionless.

In the transport layer, in the case of the usage of TCP and UDP protocols, there are a few important functionalities.
Multiplexing involves combining multiple data streams into a single transmission channel. On the other hand, de-multiplexing involves separating a single transmission channel into multiple data streams at the receiving end.

When sending messages, the TCP protocol uses four parameters: source IP address, source port number, destination IP address, and the destination port number.
When sending messages, the UDP protocol uses two parameters: the destination IP address, and the destination port number.

When using UDP, messages can get lost, or messages can be delivered out of order. Who uses UDP then? Streaming multimedia applications use UDP for the reason that they are tolerant to packet loss.

What happens if the message we sent has been altered or changed? How do we stop this?
To stop this from happening, UDP has a small algorithm inside called checksum. Checksum is an algorithm that detects flip bits. A checksum is a technique used to determine the authenticity of received data, i.e., to detect whether there was an error in transmission.