Design of Computer Networks

Mustafa Aljadery[1], Siddharth Sharma[2]

[1] *Computer Science, University of Southern California*
[2] *Computer Science, Stanford University*

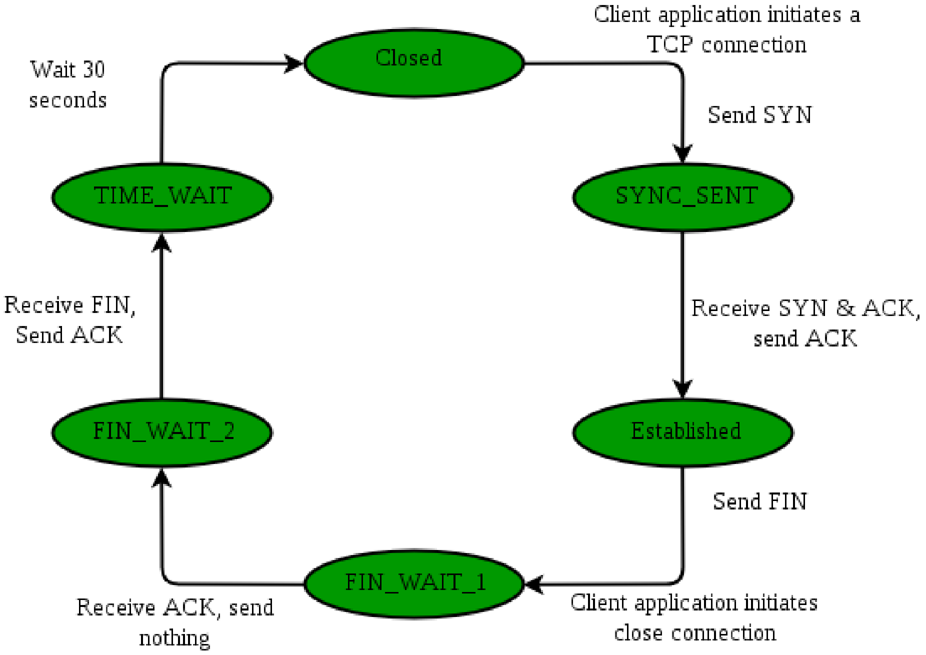
Client application initiates a TCP connection
Wait 30 seconds
Closed
Send SYN
TIME_WAIT
SYNC_SENT
Receive FIN, Send ACK
Receive SYN & ACK, send ACK
FIN_WAIT_2
Established
Send FIN
FIN_WAIT_1
Receive ACK, send nothing
Client application initiates close connection

Fig : TCP states visited by a client TCP

Contents

1 Introduction

A computer network is just a group of computers or devices connected. With a connection, they can share resources and information.

This connection of devices together is easier said than done. There are all sorts of compatibility, security, reliability, performance, and physical issues. In this textbook, we look at how we can build a computer network and address these issues.

1.1 Applications and Requirements

Applications of computer networks are vast. Anything where data can be shared between different users can be thought of as a network. The biggest example is the World Wide Web. It's literally how you are reading this textbook right now. This textbook is hosted on a server and you are getting the data from that server. Other mainstream applications include file storage and sharing, communication tools like email, gaming, and streaming media.

It's however, important to know that different people do different things on a network. If we take the Internet for example, some create applications on the Internet, those who operate and manage the network, those who design and build devices that can interact with the Internet, and more. Depending on your motivations, the applications of a computer network differ.

Going back to the motivations of a network user/operator, you must look at the incentives behind the usage. A programmer on a network would list the services they need. An operator might want the design that is easiest to manage. A designer might look at the most cost-effective design.

In general, a network must just provide connectivity among a set of computers. There only needs to be a connection of two computers together to make a computer network. These connections can be both physical, with a cable, or wireless. A gateway is a network node that connects two or more networks. It connects networks that use different protocols and architectures. Routing is selecting the best paths in a network to forward data packets towards their destination.

A requirement (or highly nice to have) of a computer network is the efficiency in moving data around. It's pretty obvious why we would want efficient moving of data, both from a resource cost-effectiveness perspective and just latency.

A strategy for cost-effective sharing is called plexing. Plexing refers to the combining of multiple signals or streams of data together over a shared medium or channel. It increases the transmission capacity of links and mediums, allows simultaneous transmission of multiple signals, and reduces costs by sharing resources effectively. Multiplexing refers to the combination of multiple input signals or data streams into one common shared link or channel for transmission. Statistical multiplexing refers to the dynamic allocation of transmission capacity between channels based on real-time traffic demand. All the resource sharing is in the form of packets. A packet is a basic unit of data transmitted over a computer network. Fairly allocating packets is the key challenge trying to be solved here.

Not only do we want our network to be cost-effective, but we also want our network to be manageable. This means we can seamlessly upgrade equipment as the network grows to carry more traffic or reach more users. There is a tradeoff here between the velocity of upgrades/downgrades and the stability of the network.

Finally, our network must be able to support a variety of devices and applications. At the end of the day, we want our network to make the lives of everyone using it easier. A channel, on a network, refers to the medium over which data is transferred from one device to another. It can be physical or wireless. There are different parts of a network that could fail and make the network unreliable, including link failures, node failures, partition failures, bit failures, and packet failures.

1.2 The Architecture of a Network

The architecture of a network is the overall design and layout of the network, which includes the components, protocols, and technologies. It is the general blueprint that designers follow. One of the best ideas when it comes to building the architecture of a network is layering. Layering refers to the process of starting with a service offered by the underlying hardware and adding layers on top, each of which provides a higher level of substructure. This is very useful because it decomposes the problem of building a network into more manageable components. Moreover, it provides a more modular design, you can decide to add or remove a service in the future without breaking the whole system, only that module.

Protocols are the rulers that define how data is transmitted and received between devices. They ensure the consistency of the communication. A service interface refers to connecting a device to a network to provide services to clients. A peer interface refers to connecting two similar devices together as peers. At the hardware level, peers directly communicate with each other over a physical medium, peer-to-peer communication is indirect. A protocol graphic is a model used to visualize the different protocols and the relationships in a layered network architecture like TCP/IP. In general, many different protocols can achieve the same task, however, there are standard bodies like IETF and ISO that establish policies for popular protocol graphs.

An important aspect of a computer network is encapsulation. Encapsulation refers to adding headers and trailers around data at each layer of the stack. It helps the protocols get the data to the right place and allows the end user to understand the data. Headers contain critical information about the message that helps both parties understand the communication message. The request header is sent from the client to the server as part of the request message. The reply header is sent from the server to the client as part of the response message.

OSI (Open System Interconnection) is a model that describes the functions of a networking system in seven layers. It provides encapsulation and abstract action from lower layers. Usually, it's used as a reference model for network protocol design. The layers are as follows:

Layer 1 - Physical Layer: Transmits raw bit streams over physical mediums.

Layer 2 - Data Link: Provides reliable transit of data frames between network nodes. Detects and corrects errors.

Layer 3 - Network Layer: Handles logical addressing and routing functions.

Layer 4 - Transport Layer: Reliable transmission of data segments between endpoints.

Layer 5 - Session Layer: Manages user sessions, dialog controls, and synchronization.

Layer 6 - Presentation Layer: Handles data formatting, compression, and encryption.

Layer 7 - Application Layer: Provides network services to applications like HTTP, SMTP, and FTP. This just allows applications to access the network

1.3 Top Layer

The success of the Internet can be interpreted as a function of the software running on general-purpose computers. It allows for an infinite number of functionalities to be run on general-purpose computers. Key aspects of computers interacting with each other are sockets and APIs. An API refers to the set of rules and protocols that allow one software application to request/receive data from another. Sockets on the other hand provide a way of communication over a network. They are equally important.

Another reason for the success of the internet is the massive increase of computational power in commodity machines like laptops. Performance can be separated into two parts, the performance of the hardware and the performance of the network as a whole. We are more focused on the latter. We want computer networks to perform well. This means we have to optimize two metrics, bandwidth and latency. Bandwidth refers to the throughput of the networks and is usually measured in bits per second that can be transmitted on the link - the logical connection between two nodes or devices in a network. Latency refers to the delay, how long it takes a message to travel from one end of the network to another.

It's important to talk about the product of both delay and bandwidth. This corresponds to how many bits the sender must transmit before the first bit arrives at the receiver. The only limitation here is the laws of physics. Furthermore, the measure of the variation in the delay of received packets is called the jitter. High jitter can negatively affect real-time applications.

2 Direct Links

A direct link is a connection between two nodes or devices that does not pass through any intermediate nodes or devices. This is the most basic computer network there is. One device is connected to another device. In general, there are many ways to connect to a protocol. You can connect to a protocol wirelessly, peer-to-peer connection, ethernet, etc.

A link in general is just a connection that allows data to be transmitted. The frequency of a link refers to the rate at which data is transmitted over a communication link. It is usually measured in hertz and indicates how many signal changes occur per second.

2.1 Encoding and Framing

The first step of passing data in a direct link is the encoding of your data format to binary data. Non-return-to-zero (NRZ) is a binary encoding scheme used in data transmission. A high or low voltage level is maintained throughout the bit period to represent binary 1 and 0 respectively. Baseline wander refers to the low-frequency variations or distortion in the baseline of a digital signal. It can occur in digital communication due to factors like cable imperfections, interference, or signal degradation.

Clock recovery is the process of extracting the timing information of a clock signal from a data stream that has been transmitted without a separate clock signal. The goal is that the sender and receiver must be synchronized in terms of timing to accurately interpret the transmitted data. Another type of encoding is Manchester encoding where each bit of data is represented by a transition from one voltage level to another within the time of a bit period.

Binary encoding in general is used for data representation where it's the language the computers understand. It's the foundation for most digital systems and technologies. On the other hand, framing is the process of dividing a stream of data into manageable frames or packets for transmission over a network. The frames usually consist of a header and a payload. The header is what contains the control information and the payload is the actual data being transmitted. Framing is essential because it allows data to be organized into discrete units that can be sent and received effectively.

Byte-oriented protocols are protocols where data is transmitted in byte-sized units. They are used in data transmission and networking. The byte is the fundamental unit of the protocol, all data is organized into bytes, and communication occurs in discrete bytes. In general, they are suitable for applications where data is naturally organized into bytes or where the byte is a convenient unit of data transmission. Usually, it's used in legacy systems.

Bit-oriented protocols are communication protocols in which data is organized and transmitted at the level of individual bits rather than bytes or larger units. These protocols are often used in low-level data transmission and communication. A bit has more fine-grained control over a byte because it's a smaller unit of representation. Given the name, the bit is the fundamental unit of a bit-oriented protocol. These protocols offer high precision and are often more efficient in terms of bandwidth utilization, as they allow for higher controls over the data being transmitted.

SONET (Synchronous Optical Networking) is a standardized protocol used in high-speed telecommunications networks. It's used for transmitting large volumes of data over optional fiber cables. The structure is to divide the data into frames, which allows for precise synchronization and error correction.

2.2 Errors and Reliability

Errors in a direct link can occur during data transmission due to factors like signal interference, noise, and imperfections in the communication medium. When errors happen, the received data may differ from the transmitted data. Reliability in direct links refers to the ability of the link to consistently and accurately transmit data with errors.

There are two types of most common errors: single-bit errors and burst errors. Single-bit errors refer to errors that occur when only one bit in the transmitted data is changed. These types of errors can be easily detected and corrected. On the other hand, burst errors are errors that involve multiple bits and occur in patterns. They are more challenging to correct and may result in data loss or corruption.

The Internet checksum algorithm is a simple error-checking algorithm used in network communication. It is employed by IP and UDP headers to detect errors in transmitted data. Cyclic Redundancy Check is a more sophisticated error-checking algorithm used in data communication, such as in-network protocols like Ethernet and data storage.

Reliability transmission is accomplished by a combination of two fundamental mechanisms. Acknowledgement and timeouts. Acknowledgment happens with. small control frame that a protocol sends back to its peer saying that it has received an earlier frame. Timeouts are when the sender does not receive an acknowledgment after a reasonable amount of time, then it retransmits the original frame.

Wait and stop is one of the checks used for reliability. After transmitting one frame, the sender waits for an acknowledgment before transmitting the next frame. If the acknowledgment does not arrive after a certain period, the sender times out and retransmits the original frame.

Another flow mechanism is a sliding window. It's better at utilization of network resources as compared to Wait and stop. It can be used in both automatic repeat request (ARQ) and selective repeat ARQ modes to handle errors and ensure reliability. It's a more efficient method for high-speed communication.

To summarize, the whole goal of reliability is to consistently deliver frames across an "unreliable" link. The sliding window algorithm can be used to reliably deliver messages over an unreliable network.

2.3 Access and Multi-Access Networks

Multi-access networks are network systems that allow multiple users or devices to access them and share resources. They're specifically designed to efficiently manage access and communication among those devices. The transceiver is the device or module that combines both the transmitting and receiving functions into a single unit. Multiple segments can be joined by repeaters.

Access protocols are a set of rules and procedures that govern how devices on a shared communication medium, such as a network or channel, can access or transmit data. These protocols are essential for efficient and fair data communication in shared environments. A MAC (Media Access Control) is a unique identifier assigned to a network interface controller of a network-connected device. A standard MAC address is 48 bits (6 bytes long), typically represented in hexadecimal format. To ensure that every adapter has a unique address, each manufacturer of ethernet devices is allocated a different predicate that must be prepended to the address on every adapter they build.

Access networks, in general, are networks that connect end-users to the broader network. They play an essential part in providing communication to the largest amount of people possible. A passive optical network uses optical fibers and passive optical components to deliver high-speed internet and other services to end-users. Network frequency bands refer to specific ranges of frequencies allocated for various wireless communication technologies. The democratization of the network edge referees to the paradigm that anyone has access to the cloud not only incumbent cloud providers or network operators.

2.4 Wireless Networks

Wireless networks are networks that use wireless communications to connect devices. They allow data to be transmitted without the need for physical cables or wired connections. The interesting thing here is the devices that use these. Most of the devices are small (most likely mobile devices), that have limited access to paper. The challenge of the network as a whole is to share the wireless medium efficiently, without interfering with each other.

Spread spectrum is a technique in wireless communication to spread the transmission of a signal over a broader frequency range than the original signal. The primary purpose of the spread spectrum is to improve the resistance to interference to provide more secure communication. Frequency hopping spread spectrum (FHSS) is an implement of spread spectrum. The carrier signal's frequency changes rapidly and periodically according to a predetermined sequence.

Wi-Fi is an example of a really popular wireless network. It allows multiple devices to connect wirelessly to a local network, typically used as a shared access point (AP). The exposed node problem occurs when a device in a Wi-Fi network refrains from transmitting data even though it could do so without causing interference. This happens when a device detects another device's transmission and incorrectly assumes that it may interfere with that transmission.

3 Connecting Networks

Connecting multiple networks is extremely important for enabling communication and data exchange between devices and users on different networks. This is also called internetworking. The overall concept is to build a large, global network.

3.1 Basics

Switching is one of the fundamental concepts of internetworking. They operate a data link layer and play a crucial role in data forwarding. They efficiently forward data to specific devices based on MAC addresses. The basic implementation is the star topology, in which all nodes are connected to a central hub or switch. The switch looks at the header of the package and sends the data to that client. To be able to identify end nodes, you have identifiers called addresses.

Datagrams are a type of data packet used to transmit information across networks, particularly in connectionless communication. Each datagram has information to enable any switch to decide how to get it to its destination. Every packet contains the complete destination address. Datagrams can be thought of as self-contained packets of data that can be independently transmitted over a packet-switched network between a source and a destination.

Source routing is when the source device sender specifies the complete route that a data packet should take through a network. It includes a list of intermediate devices that the packet should visit in sequence before reaching its destination. The great thing about this approach is it gives the sender full control over the route a packet takes through the network. It can be advantageous in situations where precise control is necessary. One of the problems with source routing is the complexity. Typically, adding source routing requires specialized configuration and support.

3.2 Ethernet

A paradigm in internetworks concerning ethernet is switch ethernet. Switched ethernet uses switches instead of hubs or repeaters to connect devices. It uses fast packet switching to provide dedicated connections and increased bandwidth, improving performance significantly compared to traditional shared Ethernet's capacity. Switched Ethernet networks are also highly scalability. Additional switches can be added to expand the network, and advanced features like VLANs can be configured for network segmentation and management.

Virtual LANs (VLANs) are a networking technology that allows you to segment a single physical network into multiple logical networks. Each VLAN operates as if it were a separate, isolated network, even though devices in different VLANs may share the same physical infrastructure. VLANs offer better network management, security, and traffic control.

The algorithm used to prevent loops in the network is the Spanning Tree Algorithm (STA). It ensures a loop-free logical topology and is particularly important in hove VLANs are implemented. This efficient loop-free and redundant logic provides great topology between bridges and switches. It's particularly useful in preventing broadcast storms and maintaining redundancy for network reliability.

3.3 Internet Protocol

The Internet is based on protocols, which are standardized systems of rules that allow computers to communicate with each other. The protocol defines how data packets should be structured and addressed so they can be transmitted across networks. IP addresses allow devices to be identified on the internet. IPv4 is the most widely used version today, while IPv6 is growing adoption.

IP is an example of an internetwork protocol - which is just an arbitrary collection of networks interconnected to provide some sort of host-ot-host packet delivery service. When we think of atheinternetwork we usually think of the Internet. IP in general is a key tool used today to build scalable, heterogeneous internetworks.

A service model refers to the type and level of service provided by a computer network or protocol. The IP service model refers to the type of service and capabilities that the Internet Protocol provides. Firstly, the service model is connectionless. It does not require a dedicated end-to-end connection like a phone circuited. Moreover, another part of the service model is that IP packets are delivered on a best-effort basis.

In general, the IP service model can be separated into an addressing scheme, which provides a way to identify all shots of an interwork and datagram model for the data delivery. The IP datagram is fundamental to the IP. A datagram is a packet sent in a connectionless manner over a network. Every diagram carries enough information to let the network forward the packer to its correct destination. This is sometimes unreliable because it does not guarantee that packets will arrive or that they will arrive in order. Packets could be lost, duplicated, delayed, or delivered out of order.

One of the main components of the IP service is the packet format. This defines the structure of data packets including header fields like source and destination IP addresses, protocol type, and more. The IP addressing service assigns logical addresses to host sand routers to identify them on the network. IPv4 and IPv6 are the addressing schemes used. Futhremreo, the IP service provides routing by determining network paths to forward packets from the source and destination. The router maintains routing tables and uses a routing algorithm to decide packet forwarding.

One of the shifts from IPv4 to IPv6 is the addressing scheme. IPv4 addresses are 32-bit binary numbers usually written in dotted decimal notation. The 32 bits are divided into two parts - the network prefix and the host identifier. The network prefix indicates the network, while the hos tID identifies a host on the network. IPv6 expands the address size to 128 bits, supporting a larger address space. In summary, the IP addressing scheme enables logical, hierarchical addressing optimized for routing packets between public and private networks across the internet.

Subnetting and classless addressing is a way to improve efficiency and flexibility compared to the original IPv4 addressing method. Subnetting involves diving a Class A, B, or C network into smaller subnetworks known as subnets. CIDR (Classless Inter-Domain Routing) enables classless prefix-based allocation for complete flexibility and efficiency in IP addressing.

Datagram forwarding is the process in which IP routes forward IP packets from one network to another to get them to their final destination. The IP routes connect two or more networks and have two or more network interfaces. Routers maintain a routing table with lists of known destination networks and the interface to use to reach them. When a router receives an incoming IP datagram, it examines the destination IP address and queries its routing table to determine which network interface leads toward that address.

Dynamic Host Configuration Protocol (DHCP) provides automated configuration of host IP addresses and other parameters on a network. It enables the host to obtain IP address and configuration settings automatically from a DHCP server. It allows devices to easily connect to the network and start communication with other hosts right away. DHCP is really important in scaling. The management of devices on a network becomes really important when you begin to scale.

To handle errors in the IP were use something called ICMP. ICMP works alongside IP to communicate issues back to the source when packets cannot be delivered or processed properly. By default, ICMP is enabled on hosts and routers. Disabling it prevents reporting, hampering diagnostics. In summary, ICMP provides a feedback and error reporting mechanism for IP. It generates informative error messages to notify sources of delivery issues so they can perform diagnostics and potentially remedy any problems.

IP tunneling is a technique to encapsulate IP packets inside another IP packet for transmission over an intermediate network. It connects two networks over an intermediate network. The intermediate network just sees the tunnel as an IP device. Tunnels can connect two remote branch networks through the public internet or connect a remote client to a private one. It's useful because it can extend private networks. It allows you to connect two remote private networks through a public network and creates a wide area network using public infrastructure.

3.4 Routing

Routing in IP refers to the process of determining paths and forwarding IP packets from a source host to a destination host across one or more IP networks. It connects two or more networks and uses routing tables to determine where to forward packets based on their destination IP address. In the most basic form, just think of it as a process for a packet to reach its intended final destination.

The Address Resolution Protocol (ARP) is used to resolve IP addresses to MAC addresses on local networks. IP addresses identify hosts on a logical network, while MAC addresses identify them at the hardware level. ARP correlates the two. When a host wants to communicate with another IP on the same network, it broadcasts an ARP request containing the destination IP. ARP bridges between layer 3 IPs and layer 2 hardware addresses.

Routing in general is a graph theory problem. The best problem is to find the lowest cost path between any two nodes, where the cost of a path equals the sum of the costs of all the edges that make up that path. The distributed nature of routing algorithms is one of the main reasons why they are a ripe area of research. That's where the complexity arises from.

The Routing Information Protocol (RIP) is a distance vector routing protocol that implements the Bellman-Forward distance vector algorithm based on the hop count as the routing metric. zit allows routes to dynamically exchange and update routing table information with neighboring routers.

OSPF (Open Shortest Path First) is a link-state routing protocol for IP networks. OSPF routes build a topological map of the entire network by flooding link-state advertisements to all routes. The routers use Dijkstra's algorithm to determine the shortest path to each destination from this topological map.

The shortest path first (SPF) algorithm computes the shortest path to each destination based on the topology and link costs of the OSPF. Large networks are divided into areas to optimize flooding and limit the size of routing tables. OSPF in general supports reqal cost multi-path routing for better load distribution and redundancy. OSPF has no limitations on hop count and metrics are based on the bandwidth of links. To summarize, OSPF provides a sophisticated link-state routing protocol that provides fast convergence, scalability, and advanced capabilities for routing within large networks.

IP, in general, provides the basic packet delivery service between networks while routing protocols determine optimal paths. Advanced protocols overcome basic limitations to enable efficient routing at the internet scale. IP routing enables the forwarding of packets between networks to reach the destination host. Protocols like ARP allow mapping between IP and hardware addresses on local networks. DHCP and NAT provide IP configuration and address-sharing capabilities. Distance vector algorithms like RIP allows routers to exchange routes with neighbors to populate routing tables.

4 Advanced Connections

In the last chapter, the biggest problem with internetwork was heterogeneity - allowing multiple different devices to access the network. Now the problem is that of scale. Scaling is essential for handling a growing number of devices on the internetwork. The idea is to help distribute traffic efficiently.

4.1 Introduction to Scale

When talking about the scalability of an internetwork, there are two things to address: scalability of routing and scalability of addressing. Both of these are crucial aspects of the IP network design.

Router scaling focuses on the capacity and performance of routers in a network. To effectively scale routers, we use high-performance hardware and implement efficient routing protocols like OSPF or BGP. Moreover, we employ load balancing to evenly distribute traffic across networks. Finally, we ensure redundancy with failover mechanisms for reliability.

A routing area is a subdivision within an OSPF routing domain. The primary and most crucial routing area in OSPF is the backbone area. It interconnects all other routing areas. All non-backbone areas must connect to the backbone area. Non-backbone areas are additional routing areas that connect to the backbone area. They help segment the OSPF network for scalability. Each non-backbone area maintains its link-state database. Routing areas in OSPF allow for more efficient routing updates and can reduce the size and complexity of the SPF tree calculation, which enhances network performance and manageability.

BGP (Border Gateway Protocol) is a protocol that helps different parts of the internetwork talk to each other. The internetwork here is an autonomous system. Autonomous systems can be thought of as systems that provide an additional way to hierarchically aggregate routing information, which allows for a lot higher scalability. Today's Internet consists of a richly interconnected set of networks, mostly operated by private companies (ISPs) rather than governments. The entire job of BGP is to prevent the establishment of looping paths. The number of nodes participating in GCP is on the order of the number of autonomous systems, which is much smaller than the number of networks. Finding a good interdomain route is only a matter of finding a path to the right border router, of which there are only a few per autonomous system.

4.2 Addressing Challenges

Another challenge with scaling is the problem of addressing space. IPv4 uses 32-bit addresses, which allow for approximately 4.3 billion unique addresses. The format is four sets of decimal numbers. The IPv6 protocol uses 128-bit addresses, which provide an almost limitless number of unique addresses. The format is written in eight sets of hexadecimal numbers. The entire motivation for defining a new version of IP is to deal with the exhaustion of the address space

IPv6 uses a classless addressing, which allows for efficient and flexible address allocation. In contrast, IPv4 uses classful addressing which divides addressing into three classes: A, B, and C. This rigid structure leads to address wastage. The IPv6 classless structure simplifies routing, reduces the need for complex subnetting, and provides a vast address space.

The packet formats of IPv4 and IPv6 are significantly different due to the distinct nature of their addressing and header structures. Packet formats are broken down into headers, header checksum, addressing, fragmentation, options, and the IP protocol field. IPv4 headers vary in length due to optional fields, commonly they are 20 bytes long without options. Ipv4 uses 32-bit source and destination addresses. Ipv5 headers are a fixed 40 bytes in length, which simplifies processing. IPv6 uses 128-bit source and destination addresses, while IPv4 uses 32-bit source and destination addresses. The Ipv4 headers include optional fields like time-to-live, type of service, and more. Ipv6 uses extension headers for optional features like fragmentation security, and routing. these headers provide flexibility. In general, Ipv6's design aims to simplify packet processing and enhance routing efficiency. It introduces larger addresses, eliminates the need for header checksums, and streamlines header structure with extension headers.

4.3 Multicast

Multicasting refers to the method of sending data from one sender to multiple recipients in a network efficiently. Unicasting on the other hand refers to one sender communicating with one receiver. Multicasting is more efficient than unicasting when sending data to multiple recipients because it reduces network traffic and minimizes the sender's workload. Data is only set to those who want it.

To better support may-to-many and one-to-many communication, IP provides an IP-level multicast analogous to the link-level multicast provided by multi-access networks like Ethernet. Distance-vector routing uses in unicast can be extended to support multicast. The resulting protocol is distance vector multicast routing protocol or DVMRP. DVMRP was the first multicast routing protocol to see widespread use.

The problem with DVMRP is that it doesn't scale well to large networks. As multicast groups and network size increase, the protocol can become inefficient in terms of resource utilization. DVMRP generates a high amount of multicast routing traffic, which can contest the network and reduce overall performance.

Before talking about newer multicast routing algorithms, we must address multicast addressing. Multicast addresses are a specific range of IP addresses used for multicast communication. They allow a single sender to send data to multiple recipients who have expressed interest in receiving the data. In IPv4, multicast addresses are in the range of 224.0.0.0n to 239.255.255.255. They are a key component of multicast communication, enabling many-to-many data transmission in IP networks.

In the simplest form, multicast routing is the process by which the multicast distribution trees are determined or, more concretely, the process by which the multicast forwarding tables are built. Distance-vector routing used in unicast can be extended to support multicast.

Interdomain Multicast (MSDP) is a protocol used in multicast routing to enable the exchange of multicast source information between autonomous systems in a network. It's a key component of multicast communication in large-scale networks, and it works in connection with other multicast routing protocols like independent multicast. MSDP allows autonomous system boundary routes to change information about multicast sources with neighboring autonomous systems. These routes use MSDP to share source information, such as the IP address of multicast sources. In general, multicast traffic can be efficiently delivered across autonomous system boundaries in large-scale networks.

Source-Speicifc Multicast (SSM) is a variation of the protocol-independent multicast routing protocol that's specifically designed for scenarios where you want to receive data from a known source. It optimizes multicast communication by allowing receivers to specify both the source and group they want to receive data from. SSm simplifies the multicast routing process because there's no need for a designated forwarder or shared trees. Receivers simply request data from specific sources and groups. Simply, SSM simplifies and optimizes multicast communication by allowing receivers to specify the exact source and group from which they want to receive data. This is valuable in applications that require efficient and controlled delivery of multicast data.

Bidirectional PIM is another multicast routing protocol, specifically designed for bidirectional communication. Unlike traditional multicast protocol, which uses shared trees or source-specific trees, BIDIR establishes a single bidirectional tree for a multicast group. BIDIR establishes a single shared bidirectional tree of each multicast group. This tree can be used for sending data from the source to the receivers and from the receivers back. However, it does require receivers to know the exact source. In general, it's a method to implement bidirectionality in multicasting.

4.4 Label Switching

Multiprotocol label switching (MPLS) efficiently directs data packets through a network base don labels. This is in contrast to traditional routing based on IP addresses. It is widely used in modern networks to improve data forwarding, quality of service, and network management.

MPLS is effective for enabling IP capabilities on devices that cannot forward IP datagrams in a normal manner. It forwards packets along explicit routes, and precalculated routes that don't necessarily match those that normal IP routing protocols would select is another benefit. Finally, it supports many types of private network services.

Destination-based forwarding is used to determine how data packets are forwarded in a network based on their destination address. It's the most common form of routing that is used by Internet protocol networks. Simply, when a router receives a packet, it looks at the destination IP address and checks its routing table to determine the next hop of interface where the packet should be sent. The packet is then forwarded to the next router or hops in the path towards its destination. This loop continues until the packet has reached the final destination. Routers make decisions on how to forward packets based on the destination IP address, ensuring they reach their intended destination efficiently and reliably.

Explicit routing is another form of routing in which the path that data packets should take through a network is explicitly defined or predetermined. The path is specified by a sequence of routers or nodes the packets should visit. It contrasts with traditional destination-based routing. Explicit routing is often used for traffic engineering purposes, where network administrators or protocols can define paths that optimize network performance, reduce congestion, or meet specific quality-of-service requirements. This is great because it gives us fine-grained control over data paths. It is commonly used in advanced networking technologies and applications to optimize network performance and resource allocation. Routers can use various algorithms to calculate explicit routes automatically.

Virtual private networks and tunnels provide secure and private communication over public or untrusted networks like the Internet. A tunnel is a secure, encrypted passage for data to travel through an untrusted network. Data sent is protected by the tunnel. Tunnels establish an end-to-end connection between two points, that ensure data integrity and privacy between those two points. VPNs, on the other hand, create a secure and encrypted connection between a user's device and a remote server or network. This ensures the data transmitted over the internet net remains private. The goal of a VPN is to mask the user's IP address, making it appear as if they are accessing the internet from a different location. VPNs use tunnels to provide secure, private, and encrypted communication over untrusted networks. VPNs establish connections and create tunnels for data to travel through ensuring privacy, security, and integrity of transmitted data, making them essential for remote access, privacy, and secure data exchange.

5 End-to-End Protocols

End-to-end protocols are protocols that provide communication between end systems or applications. They operate at the top layers of a network stack panning the the transport layer up through the application layer. They are implemented in the en dhosts but are transparent to the intervening network. They provide end-to-end functionality like ordered delivery, congestion control, and error recovery independent of the underlying network. Network routers simply forward packets and do not have to implement end-to-end protocol semantics.

At the transport layer, the common properties of a desired protocol are guarantees of message delivery, delivery of messages in the same order they are sent, delivery at most one copy of each message, support of arbitrarily large messages, support of synchronization between the sender and receiver, and the support between multiple application processes on each host.

The typical limitations of such a network are the drop in messages, the reordering of messages, the delivery of duplication copies of a given message, limited messages of finite size, and the availability to deliver messages after a long delay.

5.1 UDP

UDP (User Datagram Protocol) is the simplest possible end-to-end protocol that provides minimal datagram services. It sits above IP and provides prot numbers to enable sending and resign datagrams between multiple applications on a host. UDP has a very low overhead compared to TCP since it does not provide reliability ordering, congestion control, and recovery control. Packets may arrive out of order, duplicate, or go missing without notice. applications must handle this themselves.

There is no handshake between the source and destination before sending UDP datagrams. Datagrams can be sent at any time. Each UDP datagram contains enough information in the header to be routed to the destination without relying on prior exchanges. UDP uses a simple checksum for error detection. Datagrams that fail the checksum are silently dropped, and the underlying applications have to handle lost or corrupted data. Moreover, UDP lacks congestion control. Snders can transmit at any rate, which could contest the network. UDP is also stateless and connectionless. In general, the minimal packet overhead comes at the e cost of lacing reliability, ordering, and congestion control.

5.2 TCP

TCP is one of the core protocols of the internet and the primary transport protocol for most applications and services. It builds on top of IP by adding features like reliable data transfer, congestion control, and more. the buses are reliable, ordered, and error-checked between applications running on hosts. It uses the concept of connection between hosts, where streams of bytes are exchanged in both directions. TCP connections are established through a three-way handshake before data is transmitted. This provides connection-oriented service. Flow control mechanisms like sliding windows prevent sends from overwhelming receivers.

The format that encapsulates data is called a TCP segment. The TCP headers include the source, and destination ports, sequence and acknowledgment numbers, flags, window size, checksum, urgent points, and options. After the 20-40 byte TCP headers follow the payload data bytes being transported. The header length field specifies the total size of the header including options. Segments may be fragmented to respect IP sizes end-to-end.

The three-way handshake is the process used to establish a TCP connection between two hosts. The client sends a packet to the server with a random sequence number. this indicates the client's intention to establish a connection. The server responds with a packet that acknowledges the packet. It includes the server's sequence number and acknowledges the client's sequence number by setting the number of the first packet to +1. Finally, the client sends another packet to acknowledge to the server that the connection is established. Overall, this process enables reliable TCP connections to be established between hosts.

TCP uses sequence numbers to order bytes in the data stream between hosts. Since the sequence numbers have finite size, they can potentially wrap around and ambiguate the ordering of segments. One of the techniques to protect around this is a 32-bit sequence and acknowledgment numbers that are transferred before wrapping around. Another implementation is around a timestamp, which is the idea used in TCP.

Silly window syndrome is a phenomenon in TCP where inefficient window management leads to poor network utilization and performance. It occurs when TCP receivers advertise very small receive windows, limiting the amount of data that the sender can transmit at one time. This forces the sender to transmit tiny segments, resulting in a large number of small packets being sent back and forth. The small segments consume significant overhead in terms of headers and acknowledgments. This overhead can swamp the actual data.

Nagle's algorithm is a way to reduce small packet overhead, thus improving the network. The algorithm works by delaying sending small TCP segments if there is unacknowledged data already in flight. It will buffer small segments until an acknowledgment is received or until sufficient data is buffered to build a full-sized segment. This technique avoids sending many small packets with individual headers and contests the network.

Adaptive retransmission is used by TCP to dynamically adjust retransmission timeouts based on network conditions. The retransmission timeout determines how long TCP waits before retransmitting an unacknowledged segment. Setting the retransmission timeout requires estimating the route trip time between hosts accurately. TCP mains a retransmission timeout estimate or smoothed retransmission timeout estimate based on previous measures. The retransmission timeout is then set based on the smoothed retransmission timeout plus an additional delay to account for variance. In summary, this is a technique to optimize reliability and congestion control based on prevailing network characteristics.

The Karn/Partridge algorithm is a technique used in TCP to improve the accuracy of round-trip time estimates when packets are retransmitted. The round-trip time is measured by timing how long it takes for an acknowledgment to be received after sending a packet.t This estimate is used to calculate retransmission timeout values. Note that only retransmission timeout samples from the original non-retransmitted segments are used for calculation. This prevents retransmission from skewing the estimates.

The Jacobson/Karels algorithm is a TCZp congestion control algorithm that introduced several enhancements over the original TCP Tahoe implementation. It begins with a slow start - initially increasing the transmission rate exponentially rather than sending a full window. This prevents overwhelming the network. We then increase the transmission rate linearly rather than exponentially once the slow star finishes. This probes network capacity generally. Then we fast retransmit packets. Finally, after fast retransmit, set the congestion window to threshold plus 3packetsr rather than just 1 packet. This avoids reducing window size necessarily. Together, these improvements greatly enhance TCP's congestion control capabilities.

The two fundamental ways to measure the performance of TCP are throughput and bandwidth. Throughput is the amount of data transferred per unit of time, measured in bits/second. Indicates capacity utilized. Bandwidth is the maximum theoretical throughput available on the end-to-end path. Two other metrics that are measured are latency and jitter. Latency is the time for a segment to be sent and acknowledged, lower is better. Jitter is the variation in latency between segments. This could indicate congestion. In general, TCP congestion control algorithms like CUBIC and BBR can achieve high utilization even on high-speed long-distance links. TCP throughput degrades considerably on networks with congestion can cause some packet loss. Tuning OS and network stack parameters can optimize TCP for specific environments.

5.3 RPC

Remote Procedure Call (RPC) is a protocol that allows called procedures or functions remotely over a network. RCP provides location transparency - clients don't need to know if a procedure is local or remote. The call mechanism is the same. Moreover, RPC uses a client-server model. The client calls a procedure on a remote server as if it is local. A client sends a request message to a server, and the server responds with a reply message, with the client blocking to wait for the reply. A Transport protocol that supports the request/reply paradigm is much more than a UDP message going in one direction followed by a UDP message going in the other direction.

RPC is better thought of as a mechanism for structuring distributed systems. It's popular because it's based on the semantics of a local procedure call - the application program makes a call to a procedure without regard for whether it is local or remote. In RPC, the identifiers are used to specify the service and method you want to call. They consist of two parts the service name and the method name. The service name identifies a collection of related methods and procedures provided by a remote server. It's mainly used for a higher-level grouping of functions. You define a service with a specific name in your protocol definition. The method name identifies a specific function or procedure that you want to call within a service. You define individual methods within the service in your protocol definition. Each method has its name and parameters.

One of the biggest limits of RPC is the tight coupling between client and server, making it difficult to change the service without affecting clients. Another problem is that RPC is synchronous, which can lead to slow response times when network communication is involved. The coupling solution is to use versioning and maintain backward compatibility. Additionally, you can implement API gateways to manage changes and provide clients with stable interfaces. The solution to sync is to use asynchronous RPC implementations.

SunRPC is a protocol that enables communication between processes on a network. It was developed by Sun Microsystems and is the basis for the RPC mechanisms used in various systems. The protocol allows one program to execute procedures on a remote server as if they were local. I abstract the network communication and prove a way for processes to invoke functions on other systems. Moreover, it handles data serialization and deserialization, making it possible for processes running on different architectures and platforms to communicate. This is vital for cross-platform compatibility.

DCE-RPC is a set of RPC protocols developed by the Open Software Foundation and is designed to facilitate communication between distributed components in a networked environment. DCE-RPC allows programs running on different platforms and written in different programming languages to communicate with each other seamlessly. This is achieved through a standardized RPC process. It also includes many additional features for authentication and security.

gRPC is another RPC protocol that focuses on high performance. It's a framework that was developed by Google. It's designed for building efficient scalable distributed systems. gRPC uses protocol buffers as its interface definition language to define the structure of messages and service methods. Protobuf is language agnostic, and very efficient.t Moreover, gRPC supports a twice range of programming languages making it easy to create client and server applications that can communicate with each other.

5.4 RTP

Real-Time Transport Protocol (RTP) is used for delivering audio and video over IP networks. These are applications that typically need real-time data streaming. RTP provides the necessary mechanisms for timing, sequence numbering, and payload identification to ensure the orderly and timely delivery of multimedia data. The most basic requirement of a general-purpose multimedia protocol is that it allows similar applications to interoperate with each other.

As stated above, RTP is designed to facilitate real-time transmission of multimedia data. TRP is payload-agnostic, meaning it can carry various types of multimedia data. This allows for flexibility to adapt to different use cases. The packet header contains essential files like sequence numbers, timestamps, and the synchronization source identifier. RTP typically runs over the UDP rather than TCP to minimize latency and overhead. UDP is preferred for real-time applications despite the lack of reliability. It is also designed to promote interoperability between different technologies and platforms. This is achieved through standardized payload types and support for multiple codecs. RTP offers various profiles and extensions tailored for specific use cases.

6 Handling Scaling and Congestion

Another large problem when scaling a compute network is congestion. Congestion refers to the situation when too many packets are present in a part of a subnet. This isn't great for the network because throughput and bandwidth performance degrades and issues in reliability from packet loss increase.

A solution for congestion control is resource allocation. You can think of them as two sides of the same coin. Effective congestion control involves managing resource allocation dynamically to minimize congestion.

6.1 Resource Allocation

Resource allocation is how limited network resources are distributed among applications in the network. It allows us to efficiently handle scale. Without good resource allocation techniques, our network would become congested and affect data transfer rates and more.

A key issue with resource allocation is efficiency. Resources should be divided to maximize overall utility and satisfy the demands of a network. We want to avoid underutilized or overutilized resources. Moreover, without proper resource allocation, we would have high jitter and delay. Overall, resource allocations should scale effectively as network size, traffic, and load increase.

Resource allocation can be categorized into multiple different buckets. First, router-centric vs host-centry. Router-centric takes responsibility for deciding when packets are forwarded and selecting which packets are to be dropped. It focuses on the network itself- the routers, and links. Host-centric focuses on the end hosts of the network. These are the transport protocols and applications.

Another category is reservation-based vs feedback-based. Rservation-based allocations are resources like bandwidth which are reserved for specific flows. Reservations are made in advance based on application requirements and traffic descriptors. Once reserved, resources are guaranteed to the flow. On the other hand, feedback-based allocations are dynamically shared between flows based on feedback signals. There are no hard reservations, and allocation is adjusted based on the network state.

The last category is window-based vs rate-based. Window-based control is a process in which the sender limits transmission based on a congestion window size. That congestion window size is varied dynamically based on network feedback. Rate-base control is when the sender adjusts the transmission rate directly rather than in a window. The rate varies dynamically based on congestion feedback signals.

We also want to evaluate how our resource allocation technique is performing. The best metrics to measure are utilization, latency, and throughput. Utilization is the percentage of available bandwidth that is begin used at a given time. Latency is the time it takes for a device to receive a response after sending a request. Finally, throughput is the amount of data that can be transmitted in a given period. A good way mathematical represents our thinking is as follows:

$$Power = \frac{throughput}{delay}$$

The objective is to maximize the ratio.

6.2 Control Strategy

As we highlighted before, packets that come from multiple incoming flows are contended for limited resources in the link bandwidth. The first thing we want to do to handle those packets. To handle them we want to implement a queuing strategy. The queue determines how these resources are shared between the competing packet flows. Queue management policies control enqueue and dequeue from queues and packets drop to control congestion.

FIFO (First In First Out) is a simple queueing algorithm used in network devices, routers, and switches. Packets are queued in the order they are received in. Newly arriving packets are added to the end of the queue. This algorithm is easy to implement in hardware and software and requires minimal state to be maintained. If multiple packet flows share a FIFO queue, there is no isolation between flows. Finally, it does not differentiate between packet flows. Non-differentiation can be good or bad depending on your use case.

Fair queueing is another queue technique that ensures fair bandwidth allocation among different network flows or connections. The goal of this algorithm is to avoid a single flow from utilization of all/most of the resources, thus promoting fairness. Incoming packets are divided into separate flows based on their source, and destination. Each flow has its queue, and packets are placed in the queues.

One of the biggest challenges with queues is active queue management (AQM). AQM techniques include DECbit, RED, and ECN. DECbit (Distributed Congestion Control) is a congestion avoidance mechanism that aims to distribute congestion control responsibilities among senders and routers in a network. The routers mark packets with a DECbit based on their assessment of network congestion. The senders react to DECbit margins by reducing the sending rates to alleviate congestion. RED (Random Early Detection) monitors the queue length and randomly drops packets before the queue becomes full. The random dropping of packets serves as an early indicator of congestion to TCP senders. ECN (Explicit Congestion Notification) is a technique that uses packet marketing rather than packet dropping. Routers mark packets with an ECN field to indicate congestion. The sender responds to ECN markings by reducing its sending rate without packet loss.

Another way to control congestion is to control it from the source. Limit the rate at which data is sent from the source, such as a sender. Common source-based congestion control algorithms include Vegas, BBR, and DCTCP. Vegas is a control algorithm for TCP. Vegas continuously monitors the round-trip time and adjusts the sending rate to keep the network operating at a point where it is not congested. BBR (Bottleneck Bandwith and Round-trop Propagation Time) is a congestion algorithm that focuses on maximizing network utilization and minimizing queueing delays. It measures the bottleneck bandwidth and round-trip time to determine the optimal sending rate. DCTCP (Datacenter TCP) is optimized for data center networks. It uses ENC markings to detect and respond to congestion.

6.3 TCP

A key application of congestion control TCP. With TCP, congestion controls help regulate the flow of data over a network again preventing congestion and maintaining efficient data transfer. TCP assumes only FIFO queuing in the network's routers, but also works with fair queuing. The idea is for each source to determine how much capacity is available in the network so that it knows how many packets it can safely have in transit.

When a TCP connection is established, it brings in the slow start phase. During this phase, the sender gradually increases its sending rate to probe the available bandwidth without causing congestion. The purpose is to prevent the sender from overwhelming the network with too much data too quickly, which leads to congestion and packet loss. It's like a way to ease your way into the network. A well-implemented slow start algorithm effectively increases the congestion window exponentially, rather than linearly.

To effectively recover from packet loss and network congestion without having to wait for the normal retransmission timeout, fast retransmit and fast recovery are implemented.

These mechanisms help improve TCP in the fact of packet loss. Fast retransmit is when a sender detects that one or more of its transmitted packets are lost, it doesn't wait for the regular transmission timeout to expire, which can be relatively long. The sender retransmits the missing packet as soon as it receives three duplicate acknowledgments for the same sequence number. Fast recovery works with fast retransmit. During fast recovery, the sender reduces its congestion window size to a smaller value to reduce the rate of sending packets.

6.4 Quality of Service

Quality of Service (QoS) aims to provide different levels of service to different types of traffic, even in the presence of congestion. Several different types of algorithms can be implemented for QoS, including traffic prioritization, traffic classification, bandwidth reservation, and more. By implementing QoS, network administrators can ensure that mission-critical applications receive the necessary resources and prioritize traffic accordingly.

The first approach we are going to look at for QoS is traffic classification. Traffic classification identifies and classifies different types of network traffic and prioritizes them accordingly. It applies specified marketing or tags to those packets to indicate their priority or requirements.

Another approach is traffic prioritization. This involves preferential treatment to specific packets in a network based on important features. Each traffic class is assigned a priority level, by a numerical value of the label. High-priority traffic is given precedence over lower-priority traffic with it comes to resource allocation. A method for handling the traffic can be a queue.

Bandwidth reservation is a mechanism to guarantee a specific amount of network bandwidth for a particular traffic flow or application. It ensures that critical traffic receives the necessary resources. Network administrators can reserve a portion of the available bandwidth for specific traffic classes or applications. This reservation is often expressed as a minimum bandwidth guarantee.

Differentiated services are architectures used in computer networks to provide varying levels of service quality for different types of network traffic. They are used to classify and prioritize packets. EF (expedited forward) is used to designate high-priority traffic and is associated with real-time and latency-sensitive applications. AF(assured forwarding) classifies traffic into four different classes, each with its priority level. AF is often used for business-critical applications where varying levels of priority are needed.

7 Data Management

Both the receivers and the sender have to agree on a message format for data to be transmitted over a network. this is often called the presentation format. This data can be sent end-to-end - directly between the source and destination devices on a network, without intermediary nodes being involved. End-to-end data is complete, unprocessed data that the sending system aims to transmit to the receiving system. this is different from the packets transmitted over a network.

Multimedia over a network poses the problem of presentation and compression. Before transmitting over r a network, multimedia data is digitized and encoded into a format that is standard to the network. Compression is used to reduce the size of multimedia data for faster transmission over limited bandwidth networks. Buffering is used to temporarily store packets before presentation, to minimize jitter and ensure smooth playback. The receiver decompresses and buffers the data for high-quality, uninterrupted presentation.

7.1 Data Presentation

Presentation formatting is the process that refers to preparing data for visualization and display purposes at the receiving end of the transmission. Formatting multimedia is important because it allows us to render the data properly on the recipient's device. The challenge here is that different devices represent data in different ways.

Presentation can be categorized into three major categories base types, flat types, and complex types. Base types are the fundamental, primitive data types built into a programming language. They represent a single value. Flat types refer to data types that hold a single, simple value. They do not contain other nested values or fields. Complex types can hold multiple values and have an internal structure.

There are different strategies to convert between presentation formation. The most common strategies include standardization, intermediary formats, compressions, chucking, streaming, and more. Standardization is converting the data into a common, well-adopted data format like XML, or JSON. An intermediary format is a method for converting data to a neutral intermediary format optimized for networks, before converting again to the destination format. Compression is used libraries to convert to formats that optimize for low bandwidth networks. Chunking breaks down large data into smaller chunks during format conversion for easier transmission. Finally, streaming is the continuous conversion and transmission of data in chunks rather than the whole payload at once.

Tags are used in data presentation to annotate and categorize different elements. It's used to organize data points, which allows grouping related data points like users, products, location, and more to filter and aggregate the data. Tagging is extremely common and is really good practice.

These are common, popular network data representations and their categories. Four of the most common ones are XDR, ASN.1, NDR, ProtoBufs, and XML.

XDR is a standard data serialization format used in computer networks It allows the transfer of data between different computer caricatures. The goal is to achieve platform interoperability by defining a common wire format for data exchange between heterogeneous systems.

ASN.1 (Abstract Syntax Notation One) is a standard description language used in computer networks. It provides a formal notation to describe data structures and interactions between networked systems independent of programming languages or platforms. It defines common data types and provides ways to combine them with complex ones. In general, it's a formal abstraction for modeling data and interfaces for network-based systems and protocols.

NDR (Network Data Representation) is a framework for data serialization and RPC transport. It's Microso's equivalent of XDR and is used in their distributed computer infrastructure. The Microsoft RPC protocol uses NDR as its presentation layer to serialize method call arguments into a machine-independent format.

Protobufs (Protocol Buffers) is a language and platform-neutral data serialization format. It was developed by Google as an efficient version of CML and JSON for serializing structured data. Data structures are defined in .proto files and are compiled to generate code to target languages for serialization. Proto file definitions allow validation of encoded data. In general, it is an efficient binary serialization format for network-based messaging and data exchange between polyglot systems.

XML is a markup language used for storing and transporting data. It's a textual format for representing structured information in a standard way. XML tags allow you to enclose and organize data with custom element names, enabling semantic representation of the data. It provides a format that is machine-readable and human-readable. It's a flexible way to encapsulate and transport structured data between different systems and programs.

7.2 Multimedia

Multimedia is very useful for rich forms of communication. It allows for sending images, videos, on a computer network. As of today, it's the majority of traffic on the Internet. Part of this large shift to multimedia is the compression technology. Compress allows us to efficiently transfer multimedia across a computer network. In this section, we are going to talk about multiple compression techniques.

Lossless compression is a compression technique that reduces redundancy and entropy in the input data while allowing perfect original data reconstruction. Think of it as compression without losing any information. We want to encode a piece of data in a see of bits and encode that data in the smallest set of bits possible.

Run-length encoding is one of the simplest lossless compression techniques that work by replacing consecutive identical data values (runs of data) with a single data value and count. It works well when there are many repetitions in data. It's often used as preprocessing before applying complex compression methods.

Differential Pulse Code Modulation (DPCM) is another compression technique. It encodes the difference between the current input sample and a predicted value based on previous inputs. DPCM systems have a feedback loop where the locally decoded values are fed back to the next sample. In a sense, it is autoregressive. It's created because it efficiently removes redundancy in the input signal.

Dictionary-based compression methods work by replacing repetitions of data with references to a dictionary of data fragments seen earlier in the uncompressed data stream. It uses a dictionary of data fragments that are mined. Repeated occurrences of data fragments are replaced by a point to the location of the data fragment in the dictionary. The decoder has a copy of the dictionary, so it can reconstruct the original data using the pointers. The advantages of such a technique are that it's very adaptive and no prior knowledge of source data statistics is needed. On the other hand, it has slower compression speeds and requires more memory for holding the dictionary.

Image representations in computer networks and the data are encoded using a variety of techniques. Pixel encoding allows images to be represented as matrices of pixel intensity values. Transform coding transformed an image using Fourier, DCT, wavelets, and more.

Quantization allows image blocks to be mapped to codes from a predefined codebook. There are many more compression algorithms. The most common representations that use different compression are GIF and JPEG. GIF uses lossless compression and supports up to 8 bits per pixel. JPEG uses lossy compression, so some loss of image quality occurs to achieve a smaller file size. JPEG images can be served with different compression ratios as a quality vs file size trade-off.

For video, the most common family of compression techniques is MPEG (Moving Picture Experts Group). MPEG video compression uses both spatial and temporal redundancy reduction. Spatial redundancy is reduced using transform coding like DCT similar to JPEG images, while temporal redundancy between frames I reduced using motion estimation and compensation. MPEG uses lossy compression so some video quality I compromised to improve compression ratio.

Finally, for audio, MP3 is the most popular audio compression format. It uses lossy compression to achieve smaller file sizes for audio data. It uses perceptual coding to reduce the accuracy of less audible components of the audio to optimize for the human auditory system. Bit rates can vary from 32 kbps to 320 kbps. Higher bit rates have better quality.

8 Security

In any network, security is a critical aspect of protecting data, resources, and the integrity of the communication. Computer network security includes a wide range of measures and practices to safeguard networks from various threats and vulnerabilities. A simple technique is concealing the quantity of destination of the traffic, however, this can be taken a lot further.

One thing about security is that it will fail. In the end, security is about assuming trust, and mitigating risk. Trust and threats are two sides of the same coin. A threat is a potential failure that you design your system to avoid, and trust is an assumption you make about how external actors and internal components you build will behave.

8.1 Cryptography

Cryptographic ciphers are algorithms used to encrypt and decrypt data. Encryption transforms a message in a way that is unintelligible to any party that does now have the secret or knows how to reverse the transformation. The sender applies the encryption function to the original plain text message to result in the cipher text message. That message is then sent over the network. The receiver applies a secret decryption function - the inverse of the encryption function - to recover the original plaintext. A basic requirement for an encryption algorithm is that it can turn plain text into ciphertext in a way for the intended recipient.

Secret-key ciphers use the same cryptographic key for both encryption and decryption. The sender uses the key to encrypt the message and the receiver uses the key to decrypt the message. Common algorithms include AES, DES, Blowfish, and Twofish. What's great about secret-key ciphers is that they have the very latest encryption and are secure against brute attacks when sing sufficiently long keys. The drawback is that the sender and receiver must exchange the secret key via a secure channel. It's not suitable for an open distributed system where the sender and receiver are unknown to each other.

Public-key cryptography uses a pair of keys, a private key and a public key. The private key is used for decryption and the public key is used for encryption. We make the public key freely available but keep the private key a secret. Data encrypted with the public key can only be decrypted with the associated private key. Data encrypted with the private key can be decrypted with the public key. RSA is an example of a famous public-key cipher. The benefits of such a cipher are it allows secure communication between parties without prior exchange of keys. However, it is much slower than secret-key encryption, thus making it unsuitable for large amounts of data.

An authenticator is a credential that this stored to prove the identity of one part to another and establish trust for secure communications. This can be something like passwords, digital certificates, or tokens. Authenticators are used in many network protocols and applications including TLS, SSH, and more. Strong authentication is important to provide identities and present attacks like man-in-the-middle.

To use ciphers and authenticators, the communication participants need to know what keys to use. How does a pair of participants obtain the key to share? In general, we can break down the answer into two parts. One set of algorithms is for short-lived session keys whereas the other is for longer-lived redistributed keys. A session key is a temporary symmetric key that is generated for encrypting communications during a single session between two parties.

Public keys require each entity to have a public/private key pair. In closed systems, public keys can be manually redistributed but this affects sales. In open systems, public keys need to be distributed automatically in a secure and trusted way. Public key infrastructure (PKI) provides a hierarchical model based on certificate authorities (CAs) to bind identities to public keys via digital certificates. CAs issue digital certificates containing an entity's public key and identity. The CA digitally signs the certificate to vouch for its validity. In general, redistribution of keys allows public keys to be looked up and authenticated on first contact rather than exchanged manually.

A certification authority (CA) is an entity that issues digital certificates to validate the ownership of public keys used in public key cryptography. CAs issue digital certificates that bind a public key to an identity. The certificate is digitally signed by the CA to prove its authenticity. CAs form a hierarchical public key infrastructure (PKI). The root CA certifies intermedia CAs, which in turn issue certs to end-entities like websites. A problem with CAs is security. They can be hacked by issuing fraudulent certificates. To mitigate such risks, public CAs follow a lot of standards and audits.

Another approach to distributing and verifying public keys is a web of trust. A decentralized approach where instead of hierarchical certificate authorities verify keys, individuals sign each other's keys to establish trust. Users directly sign and certify the public keys of people they know and trust. This forms chains of trust between users. Users can incrementally build a network of trusted signatures by participating in key singing events. The advantage of such an approach is that there is no need for a centralized authority with a single point of failure, the users have more direct control. Moreover, it is resilient against compromise of any one certificate authority. On the other hand, it is hard to scale globally, and there may be no common trust anchor, the trust depends on individual circles of contact.

A problem with many of these approaches is certificate revision. We want a process of invalidating a digital certificate before it expires, usually due to security concerns. One way to do this is through browsers or operating systems. The browser can check a certificate revocation list published by this issuing CA to verify a certificate is not issues. This list is periodically issued and includes certificates revoked since the last publication date. This has its problems because of the time delay of the update, but overall it's a good step toward better security.

A problem bigger than public key redistribution is secret key redistribution. Secret key redistribution is hard because if the key is compromised in any way, the entire encryption is worthless. A couple of methods for things include manual key distribution, key enveloping, and key transport. Manual key distribution is when keys are physically delivered and installed before deployment. This doesn't scale well. Key enveloping is used to encrypt a randomly generated secret key for transmission. It provides confidentially but not authentication. Key transport is when the secret key is encrypted with the receiver's public key and signed with the sender's private key. It's one of the best methods for secret key redistribution.

Another popular approach to secret key redistribution is the Diffie-Hellman key exchange. A method that allows two parties to jointly establish a shared key over an insecure connection. The two parties first stable a secret key that can be then used for subsequent symmetric encryption of messages. Then, two numbers are agreed upon and are combined in a mathematical formula where messages can be transferred. Using this unique formula, the secret key is now shared. It's an example of a perfect forward secrecy algorithm. The only problem is its vulnerability to man-in-the-middle attacks if the identities aren't authenticated.

8.2 Authentication

We began our talk of authentication in the previous section however, authentication is a pretty complex topic when it comes to large networks. It's hard because remote entities cannot be easily identified or trusted. It's pretty hard to validate identities. Moreover, information is exposed to tampering, where packets sent over a network can be intercepted, inspected, or modified by attackers. The two biggest problems are formalized as a replay attack and the establishment of a session key.

A big problem with authentication is originality. We want to have a separate authentication mechanism for each user of the network. One way to approach this problem is to include a timestamp. The timestamp must the tamperproof so it must be covered by the authenticator. The primary drawback to timestamps is that they require distributed clock synchronization. That synchronization system itself is prone to attack. Another approach is adding a nonce. A nonce is a random number that is generated for a specific use and typically only used once. They are great because they prevent replay attacks. Each session gets a new unique nonce that cannot be reused by an attacker.

A famous public-key authentication protocol is TLS (Transport Layer Security) which uses public-key encryption to establish secure encrypted sessions between clients and servers. Servers authenticate via digital certificates signed by certification authorities. Clients can optionally authenticate using client-side certificates. It's one of the most used protocols in public-key authentication.

On the other hand, for secret-key authentication, a common protocol is Kerberos. Kerberos is a protocol that provides secure authentication for client-server applications over an insecure network. The protocol uses a third-party authentication called the key distribution center (KDC) and consists of an authentication server. The clients and servers share a secret key with the authentication server. This forms the basis of trust and mutual authentication. The client authenticates to the authentication server with the shared key. That server then returns a ticket-granting ticket for the client encrypted with the ticket-granting server. In summary, Kerberos enables authenticated and encrypted communication between any clients are servers.

8.3 Protocols and Systems

The are many systems that implement all the protocols mentioned in the previous subsections. The first system we are going to look into is Pretty Good Privacy (PGP). PGP is a system that provides cryptographic privacy and authentication for data communication. It uses a bridge cryptosystem with both symmetric-key and public-key cryptograms. The public key cryptography is used to encrypt session keys which are then used for symmetric encryption of messages. It does all of this without relying on certificate authorities.

Another system is SSH (secure shell), which is a protocol that provides secure remote access to servers and other computing devices. SSH utilizes public-key cryptography to authenticate remote machines to clients and allows encrypted connections. Remote hosts have a public/private keypair. The public key servers as the host's identity. Clients have a local copy of authorized hosts' public keys to authenticate servers they connect to. SSH establishes an encrypted tunnel protected by symmetric encryption after the initial public key authentication.

HTTPS (Hypertext Transfer Protocol Secure) is a secure version of HTTP, which is used for transmitting data over the Internet. HTTP uses TLS encryption to secure the data transmitted between the browser and the site. This is great because if someone intercepts the data, they can't read it without the encryption key. HTTPS also ensures that the data sent and received hasn't been tampered with using checksums. Digital certifications are used to prove the identities between senders and receivers.

The handshake protocol is a part of TLS. It's responsible for setting up a secure connection between a client and a server. The client begins by sending a message to the server. This message includes supported encrypted methods and other details. The server responds with its message, selecting a compatible encryption method and sending its digital certification, which includes its private key. The clients can now start a secure communication using a shared secret.

The record protocol is a protocol that is also a part of TLS and it's responsible for taking the data to be transmitted securely and turning it into secure, encrypted messages. The record protocol takes your data, locks it to make security, adds a header, sends it, and ensures that only the intended recipient can unlock and read it.

IPSec (Internet Protocol Security) is a set of protocols that are used to secure network communication at the IP layer. IPsec provides data confidentiality by encrypting the information. before sending that data, IPsec uses digital signatures and certificates to verify identities. IPsec then creates secure communication channels like tunnels.

Wireless security, 801.11i, and WPA2 are a set of protocols to secure wireless networks. WPA2 uses strong encryption, such as AES, to protect the data transmitted over the wireless network. This encryption scrambles the data, making it unreadable without the correct decryption key. It employs a robust authentication method, which verifies the identity of the users. Moreover, when a device connects to a WPA2 network, a four-way handshake occurs, which is a strong form of authentication.

9 Networking Applications

The two largest applications of computer networks are the SMTP and HTTP protocols. The SMTP protocol is used for the exchange of electronic mail. The HTTP protocol is used to communicate between web browsers and web servers. These two fall in the categories of email and web.

9.1 Email

Email has many different protocols, SMTP, MIME, IMAP, and more. SMTP is the standard protocol and sends messages between servers. SMTP handles the sending and relaying of emails from client to server and server to server. MIME is an extension of SMTP that allows non-ASCII data like images, audio, video, and more. IMAP is used for accessing email on a remote server. It allows users to download messages to a local client, keep messages on the server in mailboxes, search/select messages with filters, and manage folders/organization. The basic email message format is defined in RFC 5322. It specifies files like To, From, Subject, Date, and Body. MIME extends this to multimedia. Mail readers use IMAP to retrieve messages, interpret MIME types, display formatting, attachments, etc.

9.2 Web

The World Wide Web is an information system that runs over the internet, allowing documents and other web resources to be accessed via hyperlinks using HTTP protocols. HTTP is based on the client-server model. A client (known as a web browser) sends a request message to a server, and the server returns a response message. The request message contains a request line, headers, and an optional message body. The response message consists of a status line, response headers, and an optional message body. The status codes indicate if the request succeeded or failed. HTTP is stateless in the fact that the server doesn't remember previous requests. But cookies allow some states to be mailed across requests.

Uniform resource identifiers (URIs) are strings that identify web resources and enable browsers to retrieve them. The path points to a specific resource like a document or image. Query parameters and fragments allow for identifying portions of resources. HTPP uses TCP as its underlying transport protocol. TCP provides reliable, ordered deliver of a stream of bytes between applications. With TCP, a client must establish a connection to the server before HTTP messages can be exchanged. HTTP/3 is a new upcoming major version of HTTP after HTTP/2, currently in development. The key different in HTTP/3 is that it uses QUIC as the underlying transport protocol instead of TCP. QUIC is a transport layer network protocol designed by Google and the IETF. QUIC combines the speed benefits of UDP with the reliability of TCP along with improved security and congestion control.

Caching is one of the techniques that hold temporary storage of web documents and another resource to reduce network traffic. HTTP caching makes use of special headers like Cache-Control, Expires, and ETAg to determine if a resource is cacheable and fresh. Browsers and network caches can store a local copy of resources that can be reused without fetching it again from the server. When a browser requests a resource, the cache is checked before requesting from the origin server.

Web services are modular, self-contained applications that expose their functionality via APIs and can be accessed over the web. They allow for different software to communication with each other. Common architectures are REST and SOAP. REST (Representational State Transfer) is the most used architecture that exposes API endpoints that return JSON and XML. SOAP (Simple Object Acces Protocol) is an older XML-based messaging protocol for web services. It defines strict formats for communication.

9.3 Infrastucture and Networks

An example of the underlying infrastructure is DNS (Domain Name Service). DNS is a hierarchical decentralized dating system for computers, services, or any resource connected to the Internet. It translates domain names that are meaningful to humans into the numerical IP addresses that are used to locate and identify computer services and devices. DNS servers store DNS records that contain mappings between domain names nad IP addresses. The DNS hierarchy starts with the top-level domains and IP addresses. These domains can further be divided into subdomains. These resources are cached locally and on DNS servers to improve the performance of subsequent lookups.

A type of network that is important I san overlay network. Overlay networks provide a virtual topology that is decoupled from foursome physical topology. They're a virtual computer network that this built on top of another underlying physical network. Routers in an overlay can make forwarding decisions based on the virtual topology instead of the physical topology, which enables optimized routing.

End system miscast (ESM) is a peer-to-peer multicast architecture where the end hosts form an overlay network and cooperatively multicast data to each other, instead of relying on IP multicast at the network layer. In IP multicast, routers in the network replicate and forward multicast traffic. In ESM, end systems self-organize to distribute traffic. ESM forms an overlay mesh mong end hosts, using unicast tunnels between them to transmit multicast data over the existing IP infrastructure.

Another example of a peer-to-peer network is Gnutella. It's a decentralized network that was firs released in the early 2000s and quickly gained popularity for sharing music files without central servers. It uses a flooding search technique where peers broadcast query messages that propagate through the network until a match is found, Peers form a self-organizing overlay network by connecting randomly to find other peers. This creates a mesh topology.

One of the most popular peer-to-peer transfer protocols is BitTorrent. BitTorrent is a decentralized protocol designed for efficient file distribution and fast downloading. It uses a 'tit-for-tat' incentive mechanism where peers preferentially upload to peers who upload to them. This encourages fair sharing. Pieces are downloaded non-sequentially from multiple peers to maximize transfer speed. Rare pieces are replicated for availability.

A CDN is a distributed network of servers that provide fast delivery of internet content by caching resources closer to end users. Popular content like images, videos, web pages, and more. Wehn a user requests content, DNS resolution routes to the nearest edge server to serve the cached content quickly instead of going to the slower origin server. CDNs use load balancing and auto-scaling capabilities to handle large traffic spikes and DDoS attacks.

9.4 Blockchain

The blockchain is another example of a peer-to-peer distributed system, but instead of sharing files, it acts as a distributed ledger. The distributed ledger records transactions in a verifiable and permanent way. It is decentralized, with the ledger replicated across a peer-to-peer network of nodes. No central authority controls it. Blocks contain batches of transactions that get cryptographically linked together in a chain. Valid transactions are added to blocks which are changed together chronologically, creating an immutable record. hashing chains together blocks in a way that malicious edits or elections are easy to detect. The challenges which such a network include scalability limits, latency, energy use, and adoption barriers.

As stated above blockchain rely heavily on cryptographic hashes to secure the ledger. Hashes like SHA-256 are used extensively in Bitcoin. Transactions contain digital signatures created with the sender's private key to prove ownership. The signature is verified using the sender's public key. Consensus mechanisms like proof-of-work or proof-of-stake help prevent Sybil attacks and other ways to manipulate the ledger. Full nodes independently verify all transactions against consensus rules to maintain the high integrity of the blockchain.

In general, there are multiple promises of the blockchain. The first is decentralization. Blocks should facilitate transactions without centralized intermediaries. This promotes accessibility resilience and the removal of a single point of failure. Moreover, the blockchain should contain trustlessness, which is the ability to transact between untrusted parties in a trustless manner by using consensus rules nad cryptography. Furthermore, the blockchain should be transparent, with the visibility of all transactions to all participants. In summary, the biggest pro-blockchain argument is the case against decentralization, however, they contain many other useful properties.