Jahong Liu

System Design Basics 3: Networking Basics

Networking Basics

This is a gentle introduction to networking and maybe by seen as “networking for developers”.

We will discuss how two computers communicate with each other using the client – server architecture we discussed previously

---

What is a network?

In this example, we will be using two characters, Alice and Bob, to represent our client and server. They are widely used in the cryptographic world.

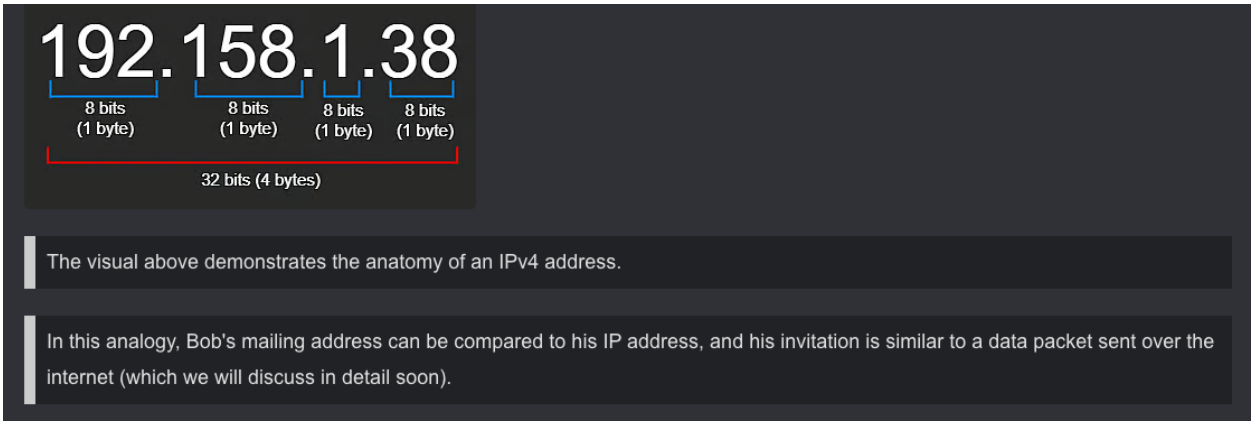
Alice and Bob are good friends living in different parts of the country. Alice wants to invite Bob to her birthday party, so she creates a personalized invitation. In networking, each device is assigned a unique IP address. This IP address acts as an identifier, allowing devices to send and receive data on the network. More specifically, Alice and Bob represent devices connected to the network.

IP Address

An IP address serves as a distinct numeric identifier for every device connected to a computer network. IP addresses come in two main types: IPv4, which is 32-bit, and IPv6, which is 128-bit. The IPv4 system theoretically allows for about 4.3 billion unique IP addresses. However due to design choices during its creation and the exponential growth of the internet the actual number of usable addresses is considerably lower. This scaricity issue led to the development and doption of IPv6, which provides a significantly larger address space to accommodate the ever-increasing number of devices on the internet.

IPv4 addresses are 32-bit and are expressed in the range 0.0.0.0- 255.255.255.255, consisting of four octets. Each octet, composed of 8 bits or a byte, can hold a maximum decimal value of 255. For further information, checking out bit manipulation from data structure and algorithms. In contrast, IPv6 addresses are 128-bit and consist of 8 octets. Each octet is represented by four hexadecimal digits and follows the format xxxx:xxxx:xxxx:xxxx:xxxx:xxxx:xxxx:xxxx. This expanded address space allows for a virtually infinite number of unique IP addresses, making address exhaustion highly improbable.

While the concepts sound promising, the practical aspect of sending the invitation from Alice to Bob needs to be addressed. This involves following specific rules known as protocols, which govern the transmission of data over the internet. One such protocol is the Internet Protocol (IP)



Protocols of sending data over a network

Let's keep following our example with Alice and Bob and see how data transfer in computer networks relates to their invitation exchange. Imagine Alice grabbing an envelope for the invitation. She records her and Bob's details on the envelope, where Bob's information on the envelope can be requested to the IP and TCP headers, and the actual invitation inside the envelope is similar to the payload or data of the packets.

IP (Internet Protocol) and data packets

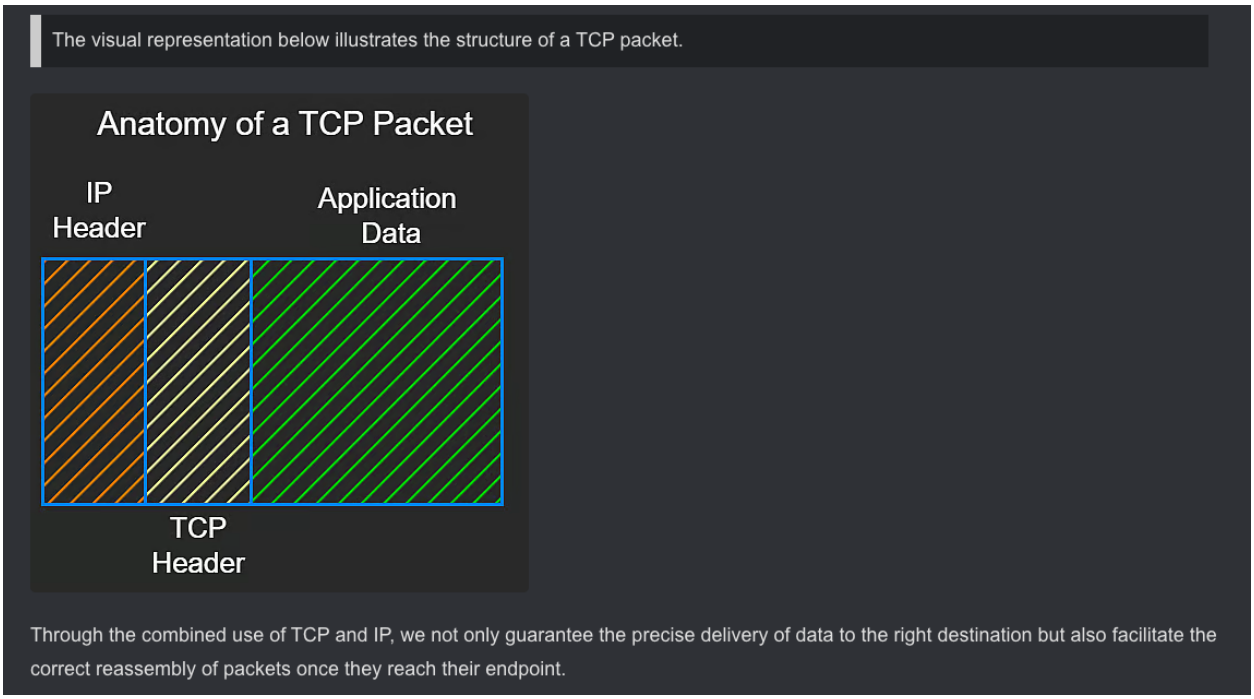
Just like Alice sends an invitation in an envelope, data is transferred over a network in the form of data packets. These packets consist of **header**, **data**(payload), and a **trailer**. The header contains important information such as the source and destination IP addresses, which are crucial for our understanding as developers.

The header is specifically referred to as the IP Header. Usually, data is divided into multiple packets, and they may not arrive in the exact order they were sent. To handle this situation, another protocol called the Transmission Control Protocol (TCP) comes into action.

TCP (Transmission Control Protocol)

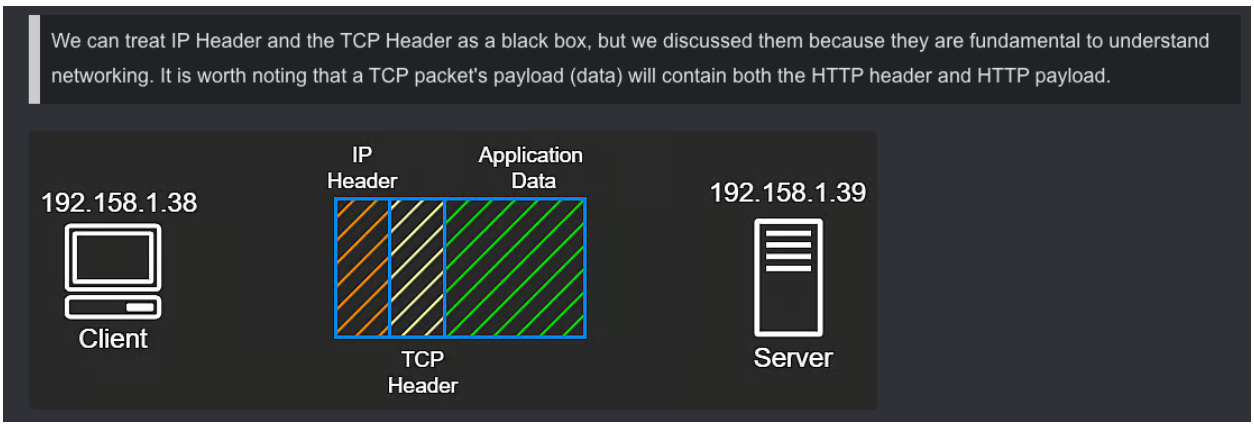
When a substantial volume of data needs to be transmitted, we often use multiple packets. TCP is responsible for the accurate transmission of these packets, ensuring they arrive in the right sequence. In a similar manner to TCP packets, consider Alice sending a multi-page letter to Bob. She assigns numbers to these pages to make sure Bob will read them in the right sequence. This means that irrespective of the order she puts the pages into the

envelope, Bob can still assemble them in the right sequence. This means that irrespective of the order she puts the pages into the envelope, Bob can still assemble them correctly upon receiving. TCP employs the same strategy by incorporating a sequence number in each packet header, making sure the packets are accurately reassembled at their destination.



Application Data

As developers, we're primarily interested in the application data portion of the packet. This data can come in different shapes, for example, an HTTP POST request, where the information we wish to transmit resides in the application segment of the packet. Conversely, it could be a GET request, whereby the data segment of the packet signifies the retrieved response.



Network Layers

Protocols within a computer network are systematically arranged into separate layers, creating a hierarchical organization. Every layer has specific duties and interacts with the layers next to it to enable communication.

For example, IP is located in the network layer, which sets the physical pathway for the transmission of data. TCP is in the transport layer, which takes care of the dependable transmission of data. HTTP operates in the application layer, making interaction between clients and servers possible.

Public vs Private Network

Another important concept is the distinction between private and public IP addresses. A public IP address is a unique identifier given to a device connected to the internet, supplied by the Internet Service Provider (ISP), and is reachable from any device on the internet. This type of IP address is utilized when communication needs to occur between devices on separate networks.

On the other hand, a private IP address is employed within a confined network, like a home or office network. Because of this, these IP addresses aren't accessible to the broader internet. As you may have noted, the main difference is the range of their accessibility. While public IP addresses can be accessed globally, private IP addresses are only reachable within a LAN (local area network).

Static vs Dynamic IP Addresses

A dynamic IP address is allocated on a temporary basis and can change every time the device establishes a connection. This type of IP address is frequently employed in home networks. They're automatically assigned, reducing the need for manual configuration.However, a drawback of dynamic IP addresses is their changeability. Hence, they're typically allocated to cliendts, not servers.

Conversely, static IP addresses require manual configuration, specifying an unchanging IP address.

> It's important to highlight that clients can also possess static IP addresses, but dynamic IP addresses are more commonly utilized for clients since they offer greater adaptability in managing the restricted pool of available IP addresses.

Ports

Application ports, also known as network ports, are numeric identifiers utilized to distinguish between multiple applications or services running on a single device. If you've ever executed a full-stack project, you'll know that you cannot operate the frontend and backend on the same port. A port number is a 16-bit integer, ranging from 0 to 65,535. Instances of these ports include port 80, typically used for HTTP. The usage of different port numbers enables the establishment of simultaneous connections.

Let's consider an Angular application as an example. By default, it runs on port 4200. We can't operate another application on the same port since it's already in use. We can visualize the concept in the diagrams below.