Networking

Contents

The Internet

Packets

Circuit switching - like old phone operators did - involves creating a dedicated communication connection between two endpoints for the duration of the data transfer. This does not work for the billions of inter-connected parts of the internet, since it is infeasible to connect them all directly. Packet switching was developed to allow a communications channel to be shared so that when one communication was not using it, another could.

When sending data across a network, data is broken into chunks called **data packets** and assembled again at the receiving end. This increases network efficiency and reliability. Each packet can hold roughly 1500 bytes. Each packet has a **TTL** (time to live). This is the number of hops before the packet is destroyed.

In packet switching, packets are sent across networks that have multiple connections with multiple routes through to a destination.

Routers forward data packets from one network to another. Each router stores data about the available routes to the destination node. It looks up the destination IP address in its routing table to find the best router to forward the packet to. Each transfer between routers is known as a hop. Routers continue to forward the packet until it reaches its destination or its TTL elapses.

At its core, a data packet is a segment of data that needs to be sent, often referred to as the **payload**. This part of the packet will vary in size from 500 to 1500 bytes. Packets also include **headers** and **trailers**.

Header
Sender’s IP address Recipient’s IP address Protocol Packet number x of y
Payload Data
Trailer End of packet flag Checksum

Packets are deliberately kept small to ensure that individual packets do not take excessive time to transfer, preventing other packets from working.

However, they should not be too small as the additional header and trailer data makes data transfer inefficient. 500-1500 bytes is an ideal compromise.

The packet header contains the recipient’s address so that it can be directed appropriately across the network. The address of the sender is also included so that replies can be sent appropriately. The packet number and overall total number of packets in the transmission is attached to enable reassembling the data. The TTL is also included.

The trailer contains an end flag, as well as error checking components that verify the data received in the payload has not been corrupted on transfer. Techniques such as checksums or CRCs are used to check the data.

The TCP/IP Protocol

A protocol is a set of rules or a formal description of a digital transmission. It will cover, for example, the contents and format of the header, the error detection, and correction procedure.

A gateway is required when data is travelling between networks that use different protocols. Networks using different transmission media can require this. Header data is removed and reapplied using the correct format of the new network. A router and gateway are often combined into one integrated device.

A protocol defines a set of rules for data communication. These must be standard across all devices, in all networks in order for communication to work. TCP/IP has become the global standard suite of networking protocols. These operate in a stack consisting of four layers.

The layers of the TCP/IP stack are as follows:

1. Application (protocol)
2. Transport (packet splitting and port numbers)
3. Internet (IP addresses)
4. Link (MAC addresses)

The computer builds the packet starting with the application layer, then the transport layer, then the internet layer, then the link layer. The router unwraps the link layer to the internet later and then adds a new link layer to send the packet between routers.

The *application layer* is used to provide services for applications to communicate across a networks, often the Internet. It uses high-level protocols that set an agreed standard between the endpoints. This could be SSH, SMTP, FTP, HTTP, HTTPS, etc. but doesn't include the port numbers. It doesn't actually determine how the data is transmitted, but rather specifies the rules of what should be sent.

The *transport layer* splits data into packets and numbers them sequentially. It adds the port number to be used based on the protocol. At the receiving end, this layer confirms that the packets have all been received and allows the receiver to request any missing packets be resent.

The *internet layer* adds the IP addresses of the sender and recipient. The router forwards each packet towards an endpoint called a socket, defined by the combination of IP address and port number. Each router uses a routing table to instruct the next hop.

The *link layer* adds the MAC address of the physical NIC that packets should be sent to based on the destination IP address. The MAC addresses change with each hop.

When the recipient receives the packet, the link layer removes the MAC address, the internet layer removes the IP address, the transport layer removes the port and reassembles the packets into the full data structure, and the application layer handles the payload itself.