Data is formatted and transmitted over the internet using various protocols and technologies. Let's break down the process into steps:

1. Data Representation: Before data is transmitted over the internet, it needs to be formatted into a suitable representation. This can include various formats like text, images, videos, audio, etc. Each type of data has its own specific format, such as ASCII or Unicode for text, JPEG or PNG for images, MP3 or WAV for audio, etc.
2. Packetization: The data is divided into smaller units called packets. Packetization breaks the data into manageable chunks for efficient transmission. Each packet typically consists of a header (containing control information) and a payload (the actual data being transmitted).
3. Addressing: Each packet needs to be addressed with the appropriate source and destination information. This is done by assigning IP addresses to devices on the network. IP (Internet Protocol) addresses uniquely identify devices and allow packets to be routed across the internet.
4. Protocol Stack: The data is then passed through a protocol stack, which is a set of networking protocols that work together to handle different aspects of data transmission. The TCP/IP (Transmission Control Protocol/Internet Protocol) protocol stack is commonly used on the internet. It includes protocols like IP, TCP, UDP (User Datagram Protocol), etc.
5. Transmission: The packets are transmitted over the physical network infrastructure, such as Ethernet or wireless networks. This transmission can happen over various mediums, including copper wires, fiber optics, or radio waves.
6. Routing: Routers on the internet examine the destination IP address in each packet's header and determine the best path for forwarding the packet towards its destination. They use routing tables to make these decisions and ensure efficient delivery.
7. Transmission Control: TCP (Transmission Control Protocol) is responsible for reliable data transmission. It ensures that packets are delivered in the correct order and handles error detection and retransmission if packets are lost or corrupted during transmission.
8. Data Link Layer: At the data link layer, the packets are encapsulated into frames that are specific to the network technology being used, such as Ethernet frames.
9. Reception and Reassembly: On the receiving end, the packets are received, and the receiving device (e.g., a computer or a server) reassembles them into the original data stream based on the packet headers and order.
10. Data Interpretation: Once the data is reassembled, it can be interpreted by the appropriate application or protocol. For example, if the data is in HTML format, a web browser will interpret the HTML tags to display the webpage.

This process of formatting, packetizing, addressing, transmitting, routing, and reassembling data allows for efficient and reliable data transfer over the internet. The specific protocols and technologies involved may vary depending on the application, network infrastructure, and communication requirements.

Difference between HTTP and HTML

HTTP (Hypertext Transfer Protocol) and HTML (Hypertext Markup Language) are two fundamental technologies that are closely related but serve different purposes on the web.

HTTP is a protocol that allows communication between web browsers (clients) and web servers. It defines how data is formatted and transmitted over the internet. When you type a website address in your browser's address bar and press Enter, your browser sends an HTTP request to the web server hosting that website. The server processes the request and sends back an HTTP response, which contains the requested data (usually HTML documents) or instructions for further actions. HTTP also supports other methods like POST, PUT, DELETE, etc., for different types of interactions with the server.

HTML, on the other hand, is a markup language used to structure and present content on the web. It provides a set of tags (enclosed in angle brackets) that define the structure and meaning of elements within a web page. HTML is responsible for defining the text, images, links, headings, paragraphs, tables, forms, and other components you see on a website. Web browsers interpret the HTML code received from the server and render it into a visual representation that users can interact with.

In summary, HTTP is the protocol that facilitates communication and data transfer between clients and servers on the web, while HTML is a markup language used to structure and present content within web pages. Both HTTP and HTML are integral parts of the World Wide Web and work together to enable the browsing and retrieval of information on websites.