

AI System Design

Mini 2: Lab 1

Marion Kipsang

mkipsang@andrew.cmu.edu

March 20, 2025

An Overview of HTTP

Hypertext Transfer Protocol (HTTP) is the backbone of web communication, enabling interactions between clients and servers in a stateless, request-response manner. It defines how requests are sent and responses are received, using various methods like GET, POST, PUT, and DELETE. Each request includes headers that provide metadata about the communication, while status codes such as 200 (OK), 404 (Not Found), and 500 (Server Error) inform the client about the result. URLs play a crucial role in identifying resources, guiding how web pages and services are accessed.

Over time, HTTP has evolved to improve performance, efficiency, and security. Early versions required new connections for each request, but HTTP/2 introduced multiplexing, persistent connections, and better compression, significantly reducing latency. HTTP/3, built on QUIC, further enhances speed and reliability by minimizing packet loss and improving encryption. The shift toward HTTPS ensures that all communication is encrypted using SSL/TLS, protecting data from interception and tampering.

Beyond web browsing, HTTP is widely used in RESTful APIs, allowing different systems to communicate over the internet. APIs leverage HTTP methods to enable CRUD (Create, Read, Update, Delete) operations, making them fundamental to web services, cloud computing, and microservices. Other important aspects include caching, authentication mechanisms, and content negotiation, all of which contribute to a more efficient and flexible web. Understanding HTTP is essential for anyone working in web development, cybersecurity, or network architecture.

What is a URL?

A Uniform Resource Locator (URL) is the structured address used to access resources on the internet. It consists of several components, including the protocol (e.g., HTTP, HTTPS, FTP), domain name (e.g., example.com), path, query parameters, and fragment identifiers. The protocol specifies how the resource should be retrieved, while the domain name identifies the web server hosting the resource. The path points to a specific file or page, and query parameters allow dynamic data to be sent in a request. Fragments refer to specific sections within a webpage, improving navigation efficiency.

URLs are essential for the functioning of the web, serving as unique identifiers for resources. They enable users to access websites, download files, and retrieve data via APIs. Proper structuring of URLs enhances SEO (Search Engine Optimization) and improves user experience. Clean, readable URLs make it easier for both users and search engines to understand the content. Additionally, relative and absolute URLs determine how resources are linked within a website, with absolute URLs specifying the full address and relative URLs pointing to locations based on the current document.

Security considerations play a crucial role in URL management. HTTPS ensures encrypted communication, protecting users from man-in-the-middle attacks and data interception. Malformed or misleading URLs, such as those used in phishing attacks, can trick users into visiting fraudulent sites. Techniques like URL encoding help handle special characters safely, preventing security vulnerabilities such as cross-site scripting (XSS) attacks. A strong understanding of URLs is essential for web developers, cybersecurity professionals, and anyone navigating the digital world.