# Recitation 2: DNS

> **Question**
>
> What is the purpose of DNS?

**Answer:** DNS operates based on Internet Protocol (IP) addresses, which serve as unique identifiers for devices connected to the internet or a local network. Websites also have corresponding IP addresses that specify the servers where they are hosted. While it is possible to access websites directly via their IP addresses (e.g., Google's IPv4 address is 8.8.8.8), memorizing these numerical sequences is impractical, especially since most people have visited several hundred (if not thousand) of websites. Instead, human-readable domain names such as "google.com" simplify navigation. The Domain Name System (DNS) functions as a naming system that maps these domain names to their respective IP addresses, enabling users to type intuitive names into their browsers rather than complex numerical addresses.

> **Question**
>
> How does DNS work?

**Answer:** DNS is structured hierarchically, with domain names segmented by periods to indicate their place within this hierarchy. For instance, in "canvas.mit.edu," the ".edu" domain represents a top-level domain (TLD), which contains the "mit" subdomain, which in turn contains "canvas."
The resolution of domain names to IP addresses involves multiple nameservers, each responsible for storing specific portions of DNS data. Every DNS query initially contacts a root nameserver, which directs the query to the relevant TLD nameserver. For example, the root nameserver knows the IP address for ".edu" (e.g., 192.14.171.191). From there, the ".edu" nameserver provides the IP address for "mit.edu" (e.g., 18.72.0.3). Finally, the "mit.edu" nameserver resolves "canvas.mit.edu" to its corresponding IP address. This hierarchical lookup process ensures efficient and organized domain name resolution.

> **Question**
>
> Why was it designed to work that way?

**Answer:** In order to understand why DNS is designed the way it is, we can explore two other possible DNS designs:

- **Centralized System**: we could have one centralized database which maintains mapping of hostnames to their corresponding I.P. addresses. There are two problems with this. First, is that if the centralized system messes up, the entire system messes up. Secondly, and more importantly, there are simply way too many requests that the centralized system would have to process. Put simply, it is infeasible.

- **Multiple Centralized Systems**: we could have multiple centralized systems too. But there's still a problem. Suppose we make an update to our centralized database–by perhaps creating a new website–we'd have to sync up all the centralized systems. Additionally, maintaining numerous centralized databases would lead to excessive and unecessary redundancy.

DNS is designed this way because, fundamentally, it is intuitive. It's easy to update, to add new websites or delete old ones. While all requests initially reach the root nameservers, these servers do not manage millions of entries directly—-and only manage their immediate subdomains.

> **Question**
>
> Who is impacted by DNS?

**Answer:** Practically everyone who uses the internet! Most people do not access websites directly via their IP addresses (I wasn't even aware that was possible until reading the textbook). There are also several companies that are DNS service providers (Google Public DNS, Cloudflare, GoDaddy, Azure DNS, and OpenDNS to name a few); they're also impacted by DNS.