Jahong Liu

System Design Basics 5: Domain Name System

Domain Name System

Domain Name System

The Domain Name System (DNS) can be likened to the internet's phone book. Just as we keep a contact list in our phones to avoid memorizing hundreds of phone numbers, DNS operates as a decentralized hierarchical naming system that converts easily readable website names into numerical IP addresses.

When we input a domain name such as google.com into our web browser, the Domin Name System (DNS) translates it into an IP address (like 142.251.211.238). This way, your computer understands where to route the request on the internet.

The overall coordination, security, and operation of domain names within DNS is managed by ICANN (Internet Corporation for Assigned Names and Numbers).

ICANN and Domain Name Registrars

The key question here is: who decides which domain name corresponds to which IP address and specifically, who it belongs to? For instance, google.com, why can't just anyone come and claim it? To answer this, we need to understand ICANN and Domain Name Registrars and how they differ. Let's consider an analogy.

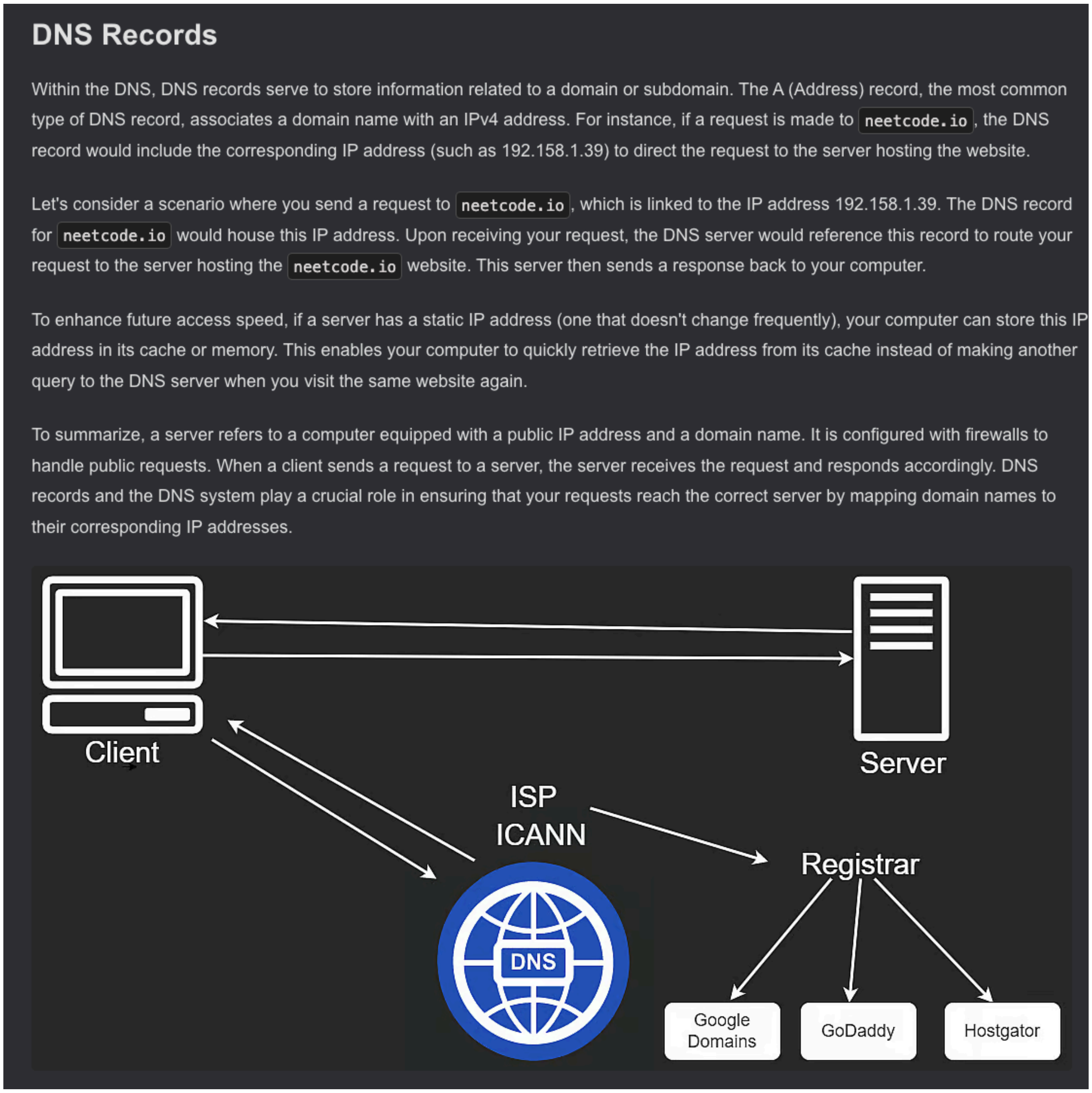
Visualize opening a store in a shopping center. ICANN could be likened to the shopping center's management company, while domain registrars resemble the individual store lease providers. Just as the shopping center's management ensures the smooth operation of the mall, ICANN ensures the efficient running of the internet's infrastructure.

When we aim to register a domain name, we engage with a domain registrar, much like how we would approach a leasing company to rent a store space in a mall. Domain registrars are certified by ICANN to offer domain name registration services. They assist us in searching for available domain names and oversee the registration procedure.

It's crucial to understand the ICAAN doesn't own domain names directly. Instead, domain names are sold by approved domain registars, such as GoDaddy, Google Domains, or HostGator, which are authorized y ICANN. These registrars maintain the registration records and ensure that the domain is registered in our name.

DNS Records

Within the DNS, DNS records serve to store information related to a domain or subdomain. The A (Address) record, the most common type of DNS record, associates a domain name with an IPv4 address. For instance, if a request is made to a website, the DNS record would include the corresponding IP address (such as 192.158.1.39) to direct the request to the server hosing the website.



DNS Records

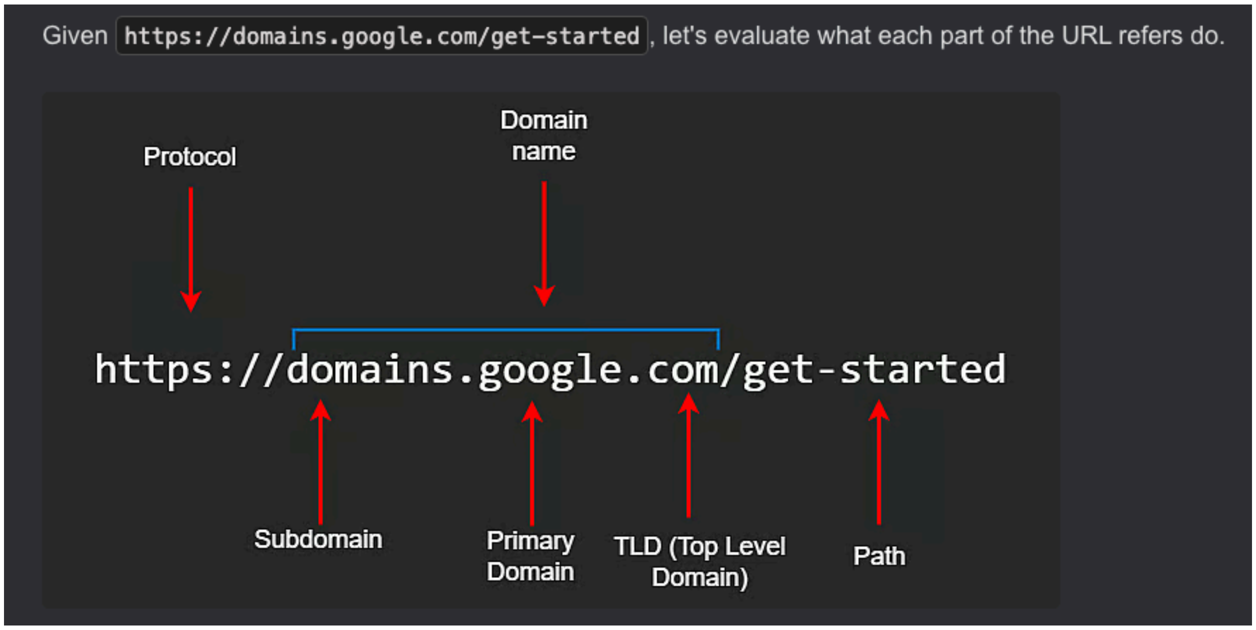
Within the DNS, DNS records serve to store information related to a domain or subdomain. The A (Address) record, the most common type of DNS record, associates a domain name with an IPv4 address. For instance, if a request is made to `neetcode.io`, the DNS record would include the corresponding IP address (such as 192.158.1.39) to direct the request to the server hosting the website.

Let's consider a scenario where you send a request to `neetcode.io`, which is linked to the IP address 192.158.1.39. The DNS record for `neetcode.io` would house this IP address. Upon receiving your request, the DNS server would reference this record to route your request to the server hosting the `neetcode.io` website. This server then sends a response back to your computer.

To enhance future access speed, if a server has a static IP address (one that doesn't change frequently), your computer can store this IP address in its cache or memory. This enables your computer to quickly retrieve the IP address from its cache instead of making another query to the DNS server when you visit the same website again.

To summarize, a server refers to a computer equipped with a public IP address and a domain name. It is configured with firewalls to handle public requests. When a client sends a request to a server, the server receives the request and responds accordingly. DNS records and the DNS system play a crucial role in ensuring that your requests reach the correct server by mapping domain names to their corresponding IP addresses.

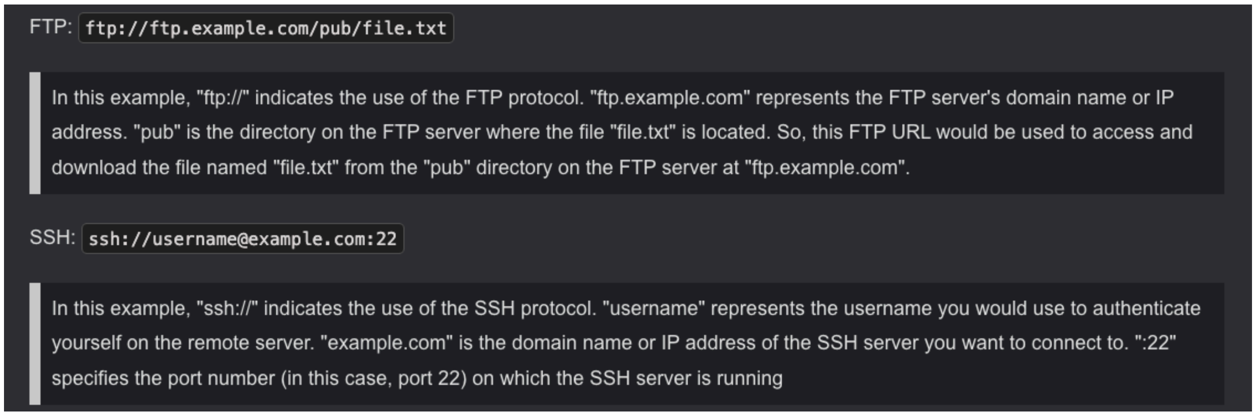
Anatomy of a URL (Uniform Resource Locator)


Given `https://domains.google.com/get-started`, let's evaluate what each part of the URL refers do.

Protocol (Scheme)

In web browsers, URLs commonly start with either HTTP or HTTPS, indicating the protocol being used. This is also commonly referred to as the scheme. It indicates which protocol to use to access the resource. HTTPS has become the more dominant protocol for URLs on the World Wide Web. However, there are several other protocols that URLs can begin with, including FTP (File Transfer Protocol) and SSH (Secure Shel).

FTP, denoted by ftp://, is utilized for accessing files and directories on remote servers. It serves as a means for transferring files between systems. On the other hand, SSH, denoted by “ssh://, is extensively employed for establishing secure remote connections to servers or computers.


FTP: `ftp://ftp.example.com/pub/file.txt`

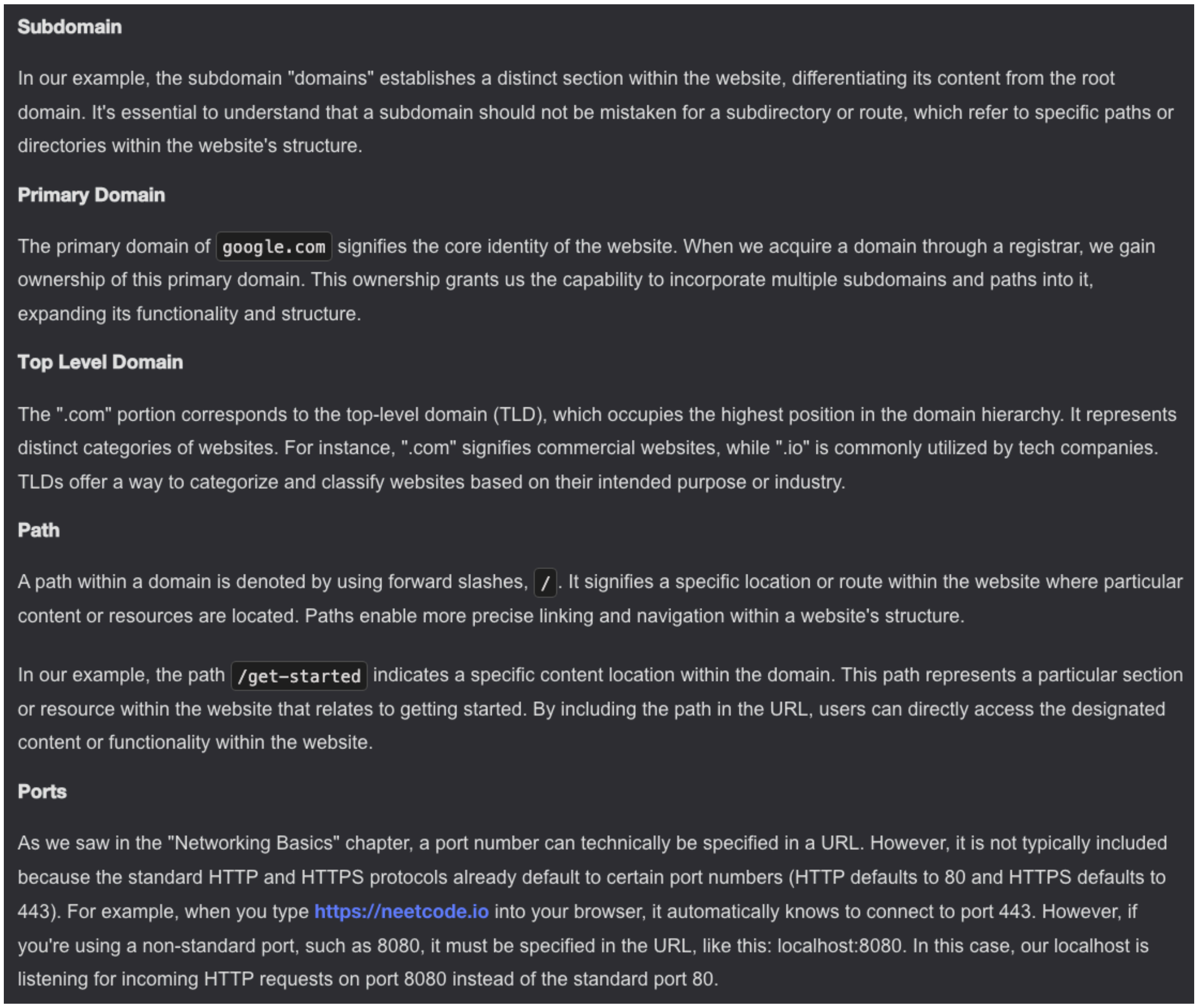
> In this example, "ftp://" indicates the use of the FTP protocol. "ftp.example.com" represents the FTP server's domain name or IP address. "pub" is the directory on the FTP server where the file "file.txt" is located. So, this FTP URL would be used to access and download the file named "file.txt" from the "pub" directory on the FTP server at "ftp.example.com".

SSH: `ssh://username@example.com:22`

> In this example, "ssh://" indicates the use of the SSH protocol. "username" represents the username you would use to authenticate yourself on the remote server. "example.com" is the domain name or IP address of the SSH server you want to connect to. ":22" specifies the port number (in this case, port 22) on which the SSH server is running

Domain

In our example, domains.google.com, the domain name can be divided into three components: the subdomain, the primary domain, and the top-level domain.



Subdomain

In our example, the subdomain "domains" establishes a distinct section within the website, differentiating its content from the root domain. It's essential to understand that a subdomain should not be mistaken for a subdirectory or route, which refer to specific paths or directories within the website's structure.

Primary Domain

The primary domain of `google.com` signifies the core identity of the website. When we acquire a domain through a registrar, we gain ownership of this primary domain. This ownership grants us the capability to incorporate multiple subdomains and paths into it, expanding its functionality and structure.

Top Level Domain

The ".com" portion corresponds to the top-level domain (TLD), which occupies the highest position in the domain hierarchy. It represents distinct categories of websites. For instance, ".com" signifies commercial websites, while ".io" is commonly utilized by tech companies. TLDs offer a way to categorize and classify websites based on their intended purpose or industry.

Path

A path within a domain is denoted by using forward slashes, `/`. It signifies a specific location or route within the website where particular content or resources are located. Paths enable more precise linking and navigation within a website's structure.

In our example, the path `/get-started` indicates a specific content location within the domain. This path represents a particular section or resource within the website that relates to getting started. By including the path in the URL, users can directly access the designated content or functionality within the website.

Ports

As we saw in the "Networking Basics" chapter, a port number can technically be specified in a URL. However, it is not typically included because the standard HTTP and HTTPS protocols already default to certain port numbers (HTTP defaults to 80 and HTTPS defaults to 443). For example, when you type https://neetcode.io into your browser, it automatically knows to connect to port 443. However, if you're using a non-standard port, such as 8080, it must be specified in the URL, like this: localhost:8080. In this case, our localhost is listening for incoming HTTP requests on port 8080 instead of the standard port 80.