ALX PROJECT
Web Infrastructure Design

Task 0.

DEFINITIONS AND EXPLANATIONS:

1. What is a server: A server is a computer program or a device that provides functionality, services, or data to other computing programs or devices. Some key points about servers:

- Servers are typically more powerful computers that are optimised to provide services and manage network resources efficiently.
- Servers operate continuously with high availability to fulfil service requests from client devices on the network.
- They allow centralised storage, management, and distribution of data, programs, communications, etc to clients.

2. What is the role of a domain name: A domain name plays several important roles:
    a. It provides an easy to remember way for humans to identify a website or other internet resource. Domain names are human readable like "www.example.com" while the machine equivalents, IP addresses, are not as friendly.
    b. It is used to route traffic to the correct web server hosting a website or internet resource. The domain name system (DNS) translates the domain name to an IP address that browsers and computers can use to find the correct server.
    c. It can be used for branding purposes and to build identity. Domain names are unique and having a brandable name for a website or online presence can contribute to better branding and marketing.

3. What type of DNS record `www` is in `www.foobar.com`: www.foobar.com contains a DNS "CNAME" record that points the subdomain "www" to the domain name foobar.com.

More specifically:

- "foobar.com" is the main domain name that resolves to the IP address of the web server hosting the website. This is an "A record".
- "www" is a common subdomain that is used to refer specifically to the website portion of the domain.

- DNS contains a "CNAME record" for the subdomain "www.foobar.com" that aliases or redirects to the main domain "foobar.com".

4. What is the role of the web server: The main roles of a web server are:
    a. Listen for incoming HTTP or HTTPS requests from clients (usually web browsers)
    b. Process requests by executing application logic to generate dynamic content
    c. Retrieve static content from storage like HTML files, images, JS files etc.
    d. Send responses back to the requesting client with the dynamic content or retrieved resources.

5. What is the role of the application server: application Server role:

- Runs the actual application code and business logic
- Performs computational tasks like algorithms, calculations
- Facilitates requesting or saving data to databases
- Generates dynamic HTML pages to display by running server-side code
- Popular environments like NodeJS, Python WSGI, Ruby on Rails
- Works together with web server like Nginx which returns output to clients

6. What is the role of the database: database role:

- Provides persistent storage and retrieval of application data
- Stores content, user data, site preferences, transactions etc
- Enables querying, filtering, analysing large datasets
- Popular database examples: MySQL, MongoDB, PostgreSQL

7. What is the server using to communicate with the computer of the user requesting the website: communication between server and user computer:

- Uses networking protocols like HTTP(S), FTP, SSH, WebSocket
- Most common is HTTP(S) which happens over port 80 or 443
- Request/response cycle - user sends HTTP request, server returns HTTP response
- Communication contents are headers, text, files, multimedia
- Encrypted with SSL/TLS certificates for secure HTTPS connections

Issues:

1. Single Point of Failure (SPOF):

- Having just one server means if that server goes down for any reason, the entire website and application will be unavailable.
- This could happen due to hardware failures, network outages, power issues, software crashes etc.
- There is no redundancy or failover if this one server fails.

2. Downtime During Maintenance:

- Operations like rebooting the server, deploying new code changes, updating the web server software or OS requires taking down the whole application leading to downtime.
- As there is no redundant server to failover to, maintenance on the single server disrupts service.

3. Scalability Limits:

- A single server has constrained hardware resources - CPU, RAM, storage and network capacity.
- As website traffic or load increases over time, at some point the hardware maxes out and gets overwhelmed.
- The application will suffer from poor performance or crashes at peak loads.
- There is no ability to scale out to additional servers to handle more demand.