



Fig 1: single server that can be reached via `www.foobar.com`

- What is a server? A **server** is a piece of computer hardware or software that provides functionality for other programs or devices, called "clients".
- What is the role of the domain name? A **domain name** is a human-readable alias for an IP address. It provides a way for people to easily remember and access websites or servers on the internet without needing to memorize the corresponding IP addresses, which are typically long strings of numbers.
- What type of DNS record `www` is in `www.foobar.com`? `www` is a **CNAME** (Canonical Name) **record**. That is, it points `www.foobar.com` to the main domain, `foobar.com`. In turn, the **A record** points `foobar.com` to the ip address `8.8.8.8`.
- What is the role of the web server? The **web server**, here Nginx, receives http requests from client computers and processes them. For requests for static content (e.g. HTML, CSS, JavaScript files, images), the web server directly serves these files to the client without involving an application server. In cases where the request requires dynamic content (e.g. data from a database, processing by an application), the web server can act as a reverse proxy. It forwards the request to an application server that can generate the dynamic content.
- What is the role of the application server? The **application server** executes backend logic, generates dynamic content, and interacts with the database to retrieve or store data for the web application.

- What is the role of the database? The **database** stores and manages the application's data, providing a persistent storage solution for information used by the application.
- What is the server using to communicate with the computer of the user requesting the website? **HTTP** (Hypertext Transfer Protocol)

Issues with the above infrastructure:

- **SPOF** (Single Point of Failure): Since there is only one server, if any hardware or software failure occurs in that server, it can't be reached.
- **Downtime when maintenance needed:** Performing maintenance tasks such as deploying new code or restarting the web server causes downtime, during which clients cannot access the server.
- **Limited Scalability:** With only one server, the infrastructure cannot easily scale to accommodate a growing user base or increased traffic. Adding more servers or resources to handle higher loads is not straightforward and may require significant redesign or downtime.