

INDIAN INSTITUTE OF TECHNOLOGY DELHI



Dev Club  
Recruitment Assignment

March 26, 2020

HARSH AGRAWAL  
2019CS50431

[Github Profile](#)

## 1 Part 1 : WARM UP

**Claim :** It would be better for the server to store the zip file of the contents of a course and provide the user the zip file instantly when the user asks for it. When new content is uploaded by a user then it is NOT necessary to update the zip file immediately, rather the admin can fix a time in a year when all zip files have to be uploaded

### **Arguments:**

1. The server experiences maximum load during the exams i.e. Minor, Majors, Quizzes etc. During these times it is highly impractical to first zip the SAME contents (since no new content is being added during exams) for each user and then provide the SAME zip file to each user.
2. For instance when Minor 2 is on the go and someone uploads the Minor 1 papers then it is not necessary to zip the Minor 1 paper of that year since the current students have already appeared for Minor 1 and they do not need it. Those who actually need it are the next yearites.
3. The admin can decide when to create the zip depending on the times when server is under the least load and most papers have been uploaded. Or better wait for BSW to collect those papers, scrap their website and make zips when students are busy enjoying their holidays!!

**Suggestion:** An approach can be to write a script that always runs on the server, it is a short script. What it will do is look for changes in the BSW website, and match it with our database. If a change is found then automatically scrap it and schedule the zip operation for sometime during the day something like lunch time 12:00 - 2:00 p.m. It will notify the admin about the scheduled operation. If admin has some objection then he can just click on a link where he can postpone the operation to some time later.

So basically this also automates the operations for admin. The script can be optimised further like scheduling the zip operations of every course into one operation rather than different operations for each paper scrapped.

## 2 Part 3 : Research

### 2.1 Networking

#### 2.1.1 What happens when we type in the URL in a web browser?

So basically when we enter something into the address bar of a web browser. It queries the DNS (Domain Name System) server for the URL of the page. What a DNS server does it actually maintains a mapping of a Domain name (human readable) into IP addresses (set of numbers that uniquely identifies a location on the internet). The DNS returns the IP to the browser. Then the browser makes a HTTP GET request to the web server hosted at that IP. The server returns a HTML response to the client (our web browser in this case). This HTML response is rendered by the web browser and displayed in a better understandable format to the user.

#### 2.1.2 Hosting a server on LAN

So python has a module named *SimpleHTTPServer* which can be used to host a server in any directory of our computer. Just type in the following command in the directory where you want to host the server

```
python3 -m SimpleHTTPServer 8080
```

This will host a server on the port 8080 in the directory where the command is run. To allow someone to visit your webpage find out your IP in the LAN network by typing *ifconfig* in the terminal and see the *IPv4* address. Suppose your IP address is 192.168.1.5 then your friend can access this webpage by typing in *http : //192.168.1.5 : 8080/* in his web browser.

**Note:** This works only for static HTML sites without any dependencies.

#### 2.1.3 Will this work on different LAN's

No, for different LAN networks like one on IITD WiFi and other on mobile data (the LAN can be the ISP's subnet), you will require a public IP address to communicate. But for this to work you will have to enable port forwarding on the IITD router which is obviously not possible. Had it been your personal router we can enable port forwarding of port 8080 to our computer.

What actually happens is when someone tries to access our public IPV4 they cannot access it directly they have to go through the router. The router performs what is called NAT (Network Address Translation) interconverting public IP to local IP. This occurs since the number of IPv4 addresses is limited and not many devices support IPv6. So instead of giving each node in a LAN a public address, only the router is given a public address and the router lends out private addresses to all the devices connected to the router.

So the router talks to the internet on each devices' behalf using different ports for each device. With port forwarding of 8080 to our private IP we can allow any external connections to the router at port 8080 to be forwarded to your computer's port 8080 on which our simple web server is hosted.

### 2.1.4 Nginx and Apache

Both are applications to host a web server on localhost. However nginx can also be used as a proxy sever. They act as web servers by constantly listening on ports for incoming connections, processes the request and responds with a reponse that the client recieves. A proxy server on the other hand makes requests to the web server on behalf of the client and returns a response to the client as if it had originally appeared from the webserver and not the proxy server.

## 2.2 Sql vs NoSql

Clearly the database is quite structured as it involves the details of a professor. The structure of the databse is unlikely to change, and to display the information quick queries are essential. Hence SQL databases seem to be the best choices. We can have different models like

- **Professor:** It will be the basic model of a professor containing the personal information of a professor.
- **Courses:** It will be a course model with details of the with Professor as a ManytoMany key.
- **Projects:** It will be a project model with details of a project with fields like Funding, Duration, Awards, Recruitment etc. It will have professor as a foreign key.

- **Research:** It will be a research table with details of the research that professor can be conducting. It can also contain fields like Description, Duration, Topic etc. It will also have professor as its foreign key.

## 2.3 Timing in JavaScript

The functions `setTimeout()` and `setInterval()` are pretty good at handling timing tasks but not all applications may be suitable for them. There are some other functions that can be used to measure timings in JavaScript as illustrated in this [GeeksforGeeks article](#)

1. Using the `performance.now()` method
2. Using the `console.time()` method
3. Using the `Date.now()` function.

However it should be noted that some exploits rely on the time taken by the server to respond and that is why JavaScript introduces some intentional delays to make very accurate time measurements practically not feasible.