# ACN Programming Assignment-2: WWW

The purpose of this www assignment was to learn about Web clients, Web servers, Web proxies, and the HyperText Transfer Protocol (HTTP).

The report is a brief explanation on 4 parts broadly:

- Part 1 is the implementation of basic web client.
- Part 2 is the implementation of basic web server.
- Part 3 is the implementation of basic web proxy.
- Part 4 is an extension to Part 1 mentioned in detail below.

## PART-1: A Simple Web Client

In this part, we develop a simple web client which will connect to the web server directly or through the intermediate web proxy using a TCP connection, send HTTP requests to the server, and display the server responses as output. **The client takes variable number of command line arguments specifying the IP address of the web proxy, host name or IP address of the web server, the port at which the proxy and web server are listening in case of indirect connection through the web proxy or simply IP address and port at which the web server is listening in case of direct connection to the web server and the path at which the requested object is stored at the web server.**

## PART-2: A Simple Web Proxy Server

In this part, we implemented and tested a simple Web proxy. This web proxy performs proxying only for HTTP requests. The goal is to build a properly functioning Web proxy for simple web pages.

In the basic operation of your proxy, it will parse, understand, and forward to the Web server, the client request. The proxy is able to handle usual response codes correctly, notifying the client as appropriate. We implemented 2 TCP (stream) sockets for client-server communication, one for client-proxy communication and the other for proxy-server communication. The proxy handles multiple clients using multi-threading.

**NOTE:** The Proxy only serves HTTP requests correctly and doesn't handle HTTPS requests aptly.

# PART-3: A Simple Web Server

In this part, we developed a multi-threaded web server that handles multiple HTTP requests simultaneously from various web client(s) (PART-1), web browsers and proxy server (PART-2). Using threading, we create a main thread in which the modified server listens for clients at a fixed port. When it receives a TCP connection request from a client, it sets up the TCP connection through another port and service the client request in a separate thread. There will be a separate TCP connection in a separate thread for each HTTP request/response pair. Each thread at the web server accepts and parses the HTTP request, gets the requested file from the server's file system, creates a HTTP response message consisting of the requested file preceded by header lines, and then sends the response directly to the client. If the requested file is not present in the server, it sends a HTTP "404 Not Found" message back to the client.

**Files in the Server directory:**

1) **index.html**
2) **style.css**

## PART-4: Extension to Proxy-Server in PART-2:

1) **English to Hindi Web PageTranslation and web usage stats at web proxy:**

   Users can type in URLs of English language web pages and get displayed web page content in Hindi language at the client side. So, Proxy accepts URL requests from the web browser, gets the response from the respective web server, translates the response message to Hindi using an open-source translation library (deep-translator) by Google, and sends the translated page to the web browser.

2) **Web usage Statistics:** The proxy also keeps track of user's browsing activities, hostnames (domains) they are visiting. Finally, the proxy provides aggregate web usage statistics on a daily basis in terms of pie charts, bar graphs.

**Scenarios and Screenshots with explanations:**

1) **Client to Local Server (Direct Connection):**
   - **Client (Local)**

```
sreyash@sreyash-VirtualBox:~/Desktop/ACN_PRG_2/Client$ python3 http_web_client.py 127.0.0.1 10000 /index.html
('127.0.0.1', 10000)
GET /index.html HTTP/1.1
Host: 127.0.0.1


HTTP/1.1 200 OK
Content-Type: text/html

<!DOCTYPE html>
<html>
<head>
    <title>Testing HTML Page with CSS</title>
    <!-- Link the external CSS file -->
    <link rel="stylesheet" type="text/css" href="style.css">
</head>
<body>
    <header>
        <h1>Welcome to My Website</h1>
    </header>
    <div>
        <p>This is a sample HTML file that uses an external CSS <span class="highlight">CSS</span> file for styling.</p>
    </div>
</body>
</html>
```
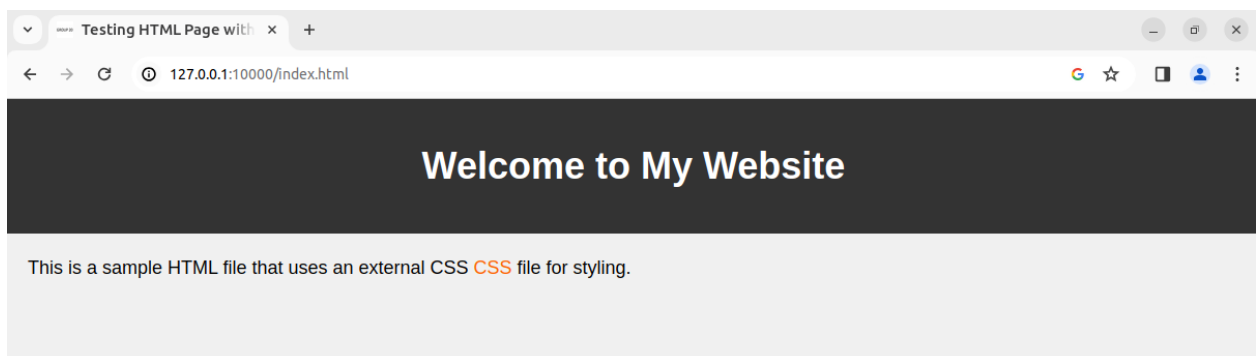
**Client**

```
sreyash@sreyash-VirtualBox:~/Desktop/ACN_PRG_2/Server$ python3 http_web_server.py
...Ready to serve...
...Client with IP address ('127.0.0.1', 47712) has connected...
...Ready to serve...
...Connection with ('127.0.0.1', 47712) served...
...Connection with ('127.0.0.1', 47712) closed...
```

**Server**

**The snippets show the case of a local client directly connecting to the local server to GET the index.html (main page) from the Server's file directory.**

- **Client (Browser)**



**Client**

**Server**

The snippets show a Browser client accessing the local server's index.html page running on port 10000.

2) **Client to Web Server (Direct Connection):**

- **Client (Local)**



**Client**

**The above snippets show the local client retrieving the web page from the http web server www.example.com running on port 80.**
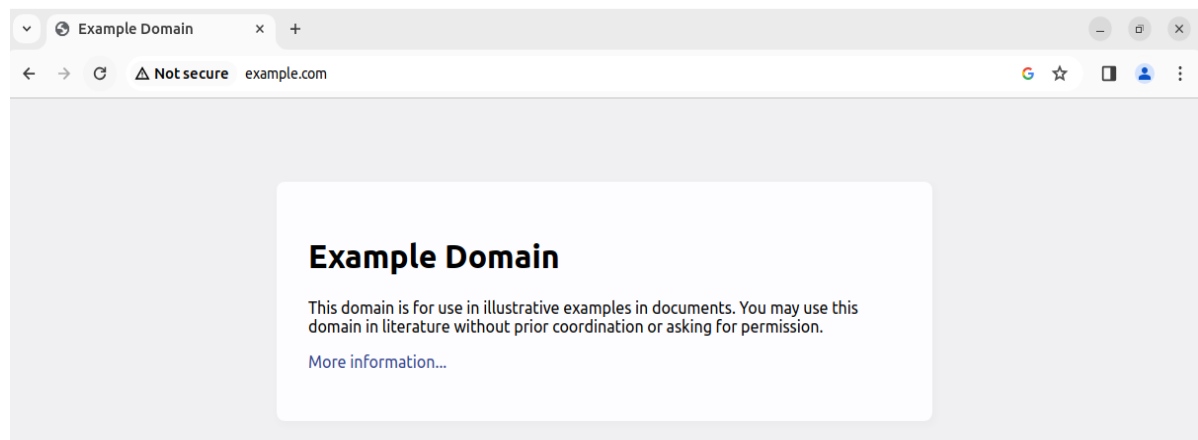
**3) Client to Web Server (Indirect Connection - Through Proxy)**

- **Client (Browser)**



```
sreyash@sreyash-VirtualBox:~/Desktop/ACN_PRG_2/Server$ python3 http_proxy_server.py
...Proxy Server is ready to serve...
Received a connection from:  ('127.0.0.1', 49716)
...Proxy Server is ready to serve...
Received a connection from:  ('127.0.0.1', 49732)
...Proxy Server is ready to serve...
Received request from a client ('127.0.0.1', 49716):  GET http://www.example.com/ HTTP/1.1
Host: www.example.com
Proxy-Connection: keep-alive
Cache-Control: max-age=0
Upgrade-Insecure-Requests: 1
User-Agent: Mozilla/5.0 (X11; Linux x86_64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/118.0.0.0 Safari/537.36
Accept: text/html,application/xhtml+xml,application/xml;q=0.9,image/avif,image/webp,image/apng,*/*;q=0.8,application/signed-exchange
;v=b3;q=0.7
Accept-Encoding: gzip, deflate
Accept-Language: en-GB,en-US;q=0.9,en;q=0.8


...Sending request to the main web server...
^C
Connection Terminated
```

**Proxy Server**



**Client**

**The above snippets show an indirect connection sent from a client through the proxy server to www.example.com and the browser displays the page.**

- **Client (Local)**

```
sreyash@sreyash-VirtualBox:~/Desktop/ACN_PRG_2/Client$ python3 http_web_client.py www.example.com 80 /index.html 35000 127.0.0.1
('127.0.0.1', 35000)
GET /index.html HTTP/1.1
Host: www.example.com:80


HTTP/1.1 200 OK
Accept-Ranges: bytes
Age: 500265
Cache-Control: max-age=604800
Content-Type: text/html; charset=UTF-8
Date: Sun, 05 Nov 2023 12:02:58 GMT
Etag: "3147526947+gzip"
Expires: Sun, 12 Nov 2023 12:02:58 GMT
Last-Modified: Thu, 17 Oct 2019 07:18:26 GMT
Server: ECS (dcb/7F83)
Vary: Accept-Encoding
X-Cache: HIT
Content-Length: 1256

<!doctype html>
<html>
<head>
    <title>Example Domain</title>

    <meta charset="utf-8" />
    <meta http-equiv="Content-type" content="text/html; charset=utf-8" />
    <meta name="viewport" content="width=device-width, initial-scale=1" />
    <style type="text/css">
    body {
        background-color: #f0f0f2;
        margin: 0;
        padding: 0;
        font-family: -apple-system, system-ui, BlinkMacSystemFont, "Segoe UI", "Open Sans", "Helvetica Neue", Helvetica, Arial, sans
-serif;
```

```
    }
    div {
        width: 600px;
        margin: 5em auto;
        padding: 2em;
        background-color: #fdfdff;
        border-radius: 0.5em;
        box-shadow: 2px 3px 7px 2px rgba(0,0,0,0.02);
    }
    a:link, a:visited {
        color: #38488f;
        text-decoration: none;
    }
    @media (max-width: 700px) {
        div {
            margin: 0 auto;
            width: auto;
        }
    }
    </style>
</head>

<body>
<div>
    <h1>Example Domain</h1>
    <p>This domain is for use in illustrative examples in documents. You may use this
    domain in literature without prior coordination or asking for permission.</p>
    <p><a href="https://www.iana.org/domains/example">More information...</a></p>
</div>
</body>
</html>
```

**Client**

```
sreyash@sreyash-VirtualBox:~/Desktop/ACN_PRG_2/Server$ python3 http_proxy_server.py
...Proxy Server is ready to serve...
Received a connection from:  ('127.0.0.1', 56098)
...Proxy Server is ready to serve...
Received request from a client ('127.0.0.1', 56098):  GET /index.html HTTP/1.1
Host: www.example.com:80


...Sending request to the main web server...
^C
Connection Terminated
```

**Proxy Server**

**In the above snippets the local web client sends its GET request through the proxy to get the response from www.example.com running on port 80.**

### 4) Client to Local Server (Indirect Connection - Through Proxy)

- **Client (Local)**

```
sreyash@sreyash-VirtualBox:~/Desktop/ACN_PRG_2/Client$ python3 http_web_client.py 127.0.0.1 10000 /index.html 35000 127.0.0.1
('127.0.0.1', 35000)
GET /index.html HTTP/1.1
Host: 127.0.0.1:10000


HTTP/1.1 200 OK
Content-Type: text/html

<!DOCTYPE html>
<html>
<head>
    <title>Testing HTML Page with CSS</title>
    <!-- Link the external CSS file -->
    <link rel="stylesheet" type="text/css" href="style.css">
</head>
<body>
    <header>
        <h1>Welcome to My Website</h1>
    </header>
    <div>
        <p>This is a sample HTML file that uses an external CSS <span class="highlight">CSS</span> file for styling.</p>
    </div>
</body>
</html>
```

**Client**

```
sreyash@sreyash-VirtualBox:~/Desktop/ACN_PRG_2/Server$ python3 http_proxy_server.py
...Proxy Server is ready to serve...
Received a connection from:  ('127.0.0.1', 58418)
...Proxy Server is ready to serve...
Received request from a client ('127.0.0.1', 58418):  GET /index.html HTTP/1.1
Host: 127.0.0.1:10000


...Sending request to the main web server...
Received a connection from:  ('127.0.0.1', 49060)
...Proxy Server is ready to serve...
Received request from a client ('127.0.0.1', 49060):  GET /index.html HTTP/1.1
Host: 127.0.0.1:10000


...Sending request to the main web server...
Received a connection from:  ('127.0.0.1', 57292)
...Proxy Server is ready to serve...
Received request from a client ('127.0.0.1', 57292):  GET /index.html HTTP/1.1
Host: 127.0.0.1:10000
```

-
**Proxy Server**

```
sreyash@sreyash-VirtualBox:~/Desktop/ACN_PRG_2/Server$ python3 http_web_server.py
...Ready to serve...
...Client with IP address ('127.0.0.1', 37446) has connected...
...Connection with ('127.0.0.1', 37446) served...
...Connection with ('127.0.0.1', 37446) closed...
...Ready to serve...
...Client with IP address ('127.0.0.1', 35318) has connected...
...Ready to serve...
...Connection with ('127.0.0.1', 35318) served...
...Connection with ('127.0.0.1', 35318) closed...
```

**Server**

In the above 3 snippets the local web client sends its GET request through the Proxy to the local web server running on port 10000 and is served with the index.html file.

## PART - 4:

- **English to Hindi Translation:**

```
<!DOCTYPE html>

<html>
<head>
<title>Example Domain</title>
<meta charset="utf-8"/>
<meta content="text/html; charset=utf-8" http-equiv="Content-type"/>
<meta content="width=device-width, initial-scale=1" name="viewport"/>
<style type="text/css">
    body {
        background-color: #f0f0f2;
        margin: 0;
        padding: 0;
        font-family: -apple-system, system-ui, BlinkMacSystemFont, "Segoe UI", "Open Sans", "Helvetica Neue", Helvetica, Arial, sans
-serif;

    }
    div {
        width: 600px;
        margin: 5em auto;
        padding: 2em;
        background-color: #fdfdff;
        border-radius: 0.5em;
        box-shadow: 2px 3px 7px 2px rgba(0,0,0,0.02);
    }
    a:link, a:visited {
        color: #38488f;
        text-decoration: none;
    }
    @media (max-width: 700px) {
        div {
            margin: 0 auto;
            width: auto;
        }
    }
    </style>
```

```
</head>
<body>
<div>
<h1>उदा हरण डोमेन</h1>
<p>यह डोमेन दस्तावेज़ों में उदाहरणात्मक उदाहरणों में उपयोग के लिए है। अ इस्का प्रयोग कर सक्ते हैं
    पूर्व समन्वय या अनुमति मांगे बिना साहित्य में डोमेन।</p>
<p>अधिक जानकारी ...</p>
</div>
</body>
</html>
```
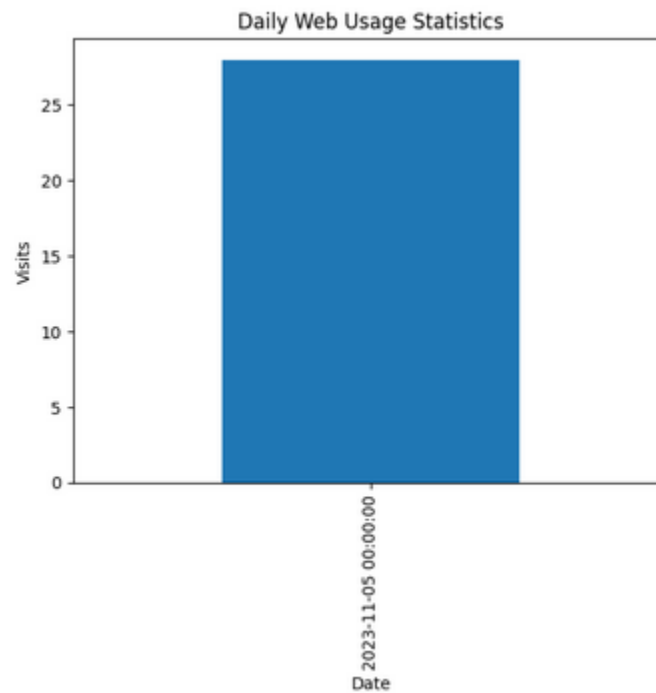
**Client**

In the above snippets it is seen that the web page [www.example.com](www.example.com) requested by the local web client which is sent through the proxy to the main web server is returned to the client, translated in Hindi. The proxy parses the response from the server and modifies the html file by translating it and sending it back to the client.
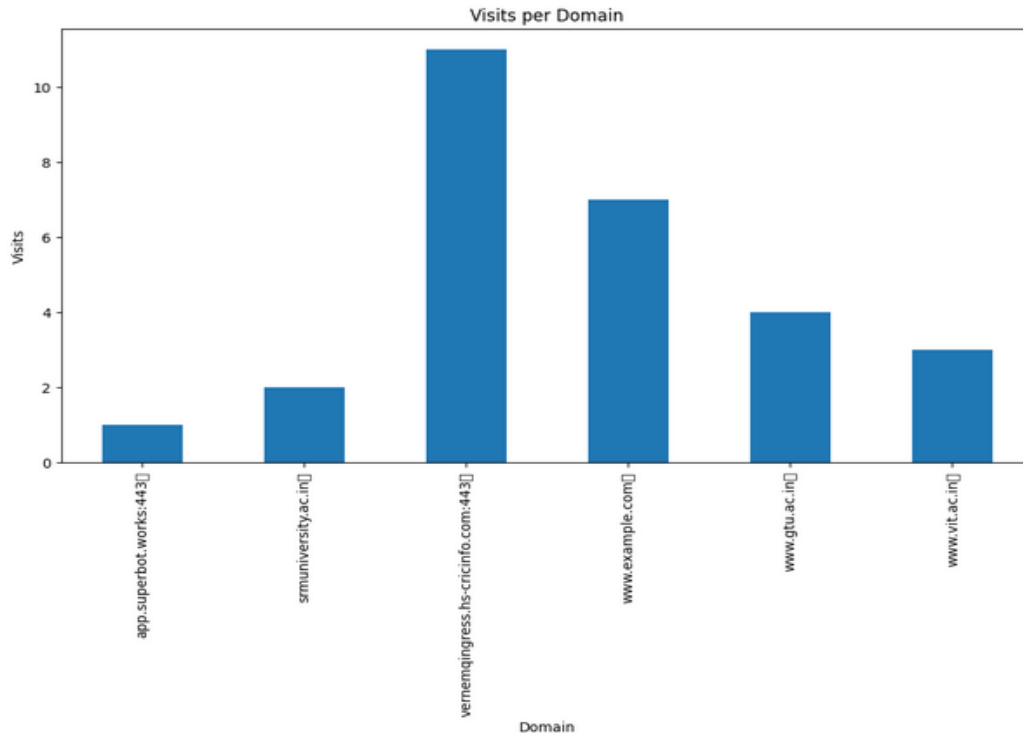
## - Web-usage Statistics:



**Figure 1. The total number of domains visited / day**
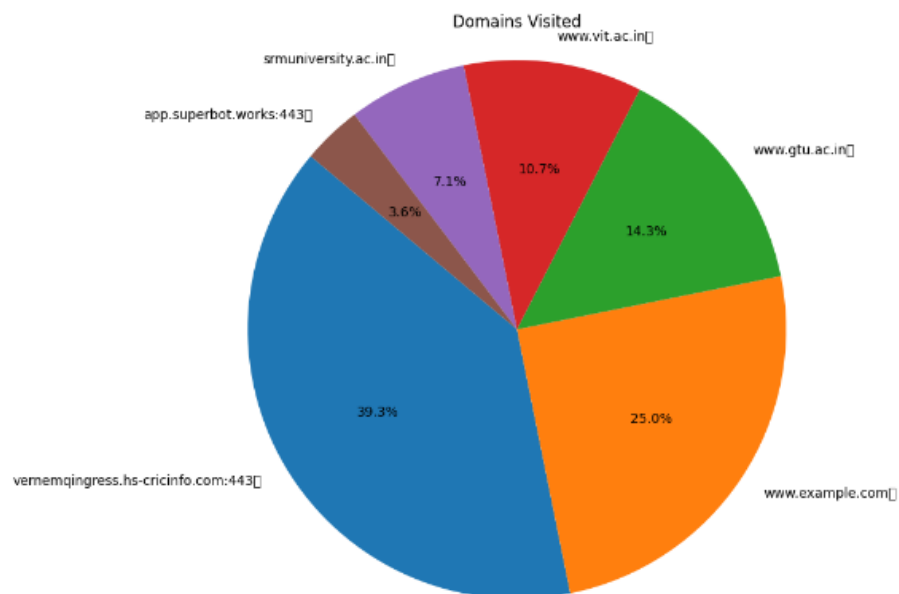


**Figure 2. Pie Chart for the count of no. of different domains seen to be visited over the session**

**Figure 3. Bar Chart for the count of no. of different domains seen to be visited over the session**

# ANTI-PLAGIARISM Statement

*We certify that this assignment/report is our own work, based on our personal study and/or research and that we have acknowledged all material and sources used in its preparation, whether they be books, articles, packages, datasets, reports, lecture notes, and any other kind of document, electronic or personal communication. We also certify that this assignment/report has not previously been submitted for assessment/project in any other course lab, except where specific permission has been granted from all course instructors involved, or at any other time in this course, and that we have not copied in part or whole or otherwise plagiarized the work of other students and/or persons. Additionally, we acknowledge that we may have used AI tools, such as language models (e.g., ChatGPT, Bard), for assistance in generating and refining my assignment, and we have made all reasonable efforts to ensure that such usage complies with the academic integrity policies set for the course. I pledge to uphold the principles of honesty and responsibility at CSE@IITH. In addition, we understand our responsibility to report honour violations by other students if we become aware of it.*

Names <Roll Nos>: **Sreyash Mohanty CS23MTECH14015**
**Rohit Sutrave  CS23MTECH14010**
**Rathwa Yuvrajsinh Jayantibhai SM23MTECH11011**

Date: **05/11/2023**

Signatures:  **SREYASH MOHANTY**
             **ROHIT SUTRAVE**
             **RATHWA YUVRAJSINH JAYANTIBHAI**