



# **BÁO CÁO ĐỒ ÁN PROXY SERVER**

*TRƯỜNG ĐH KHTN TP HCM  
KHOA CÔNG NGHỆ THÔNG TIN  
MÔN HỌC: MẠNG MÁY TÍNH*

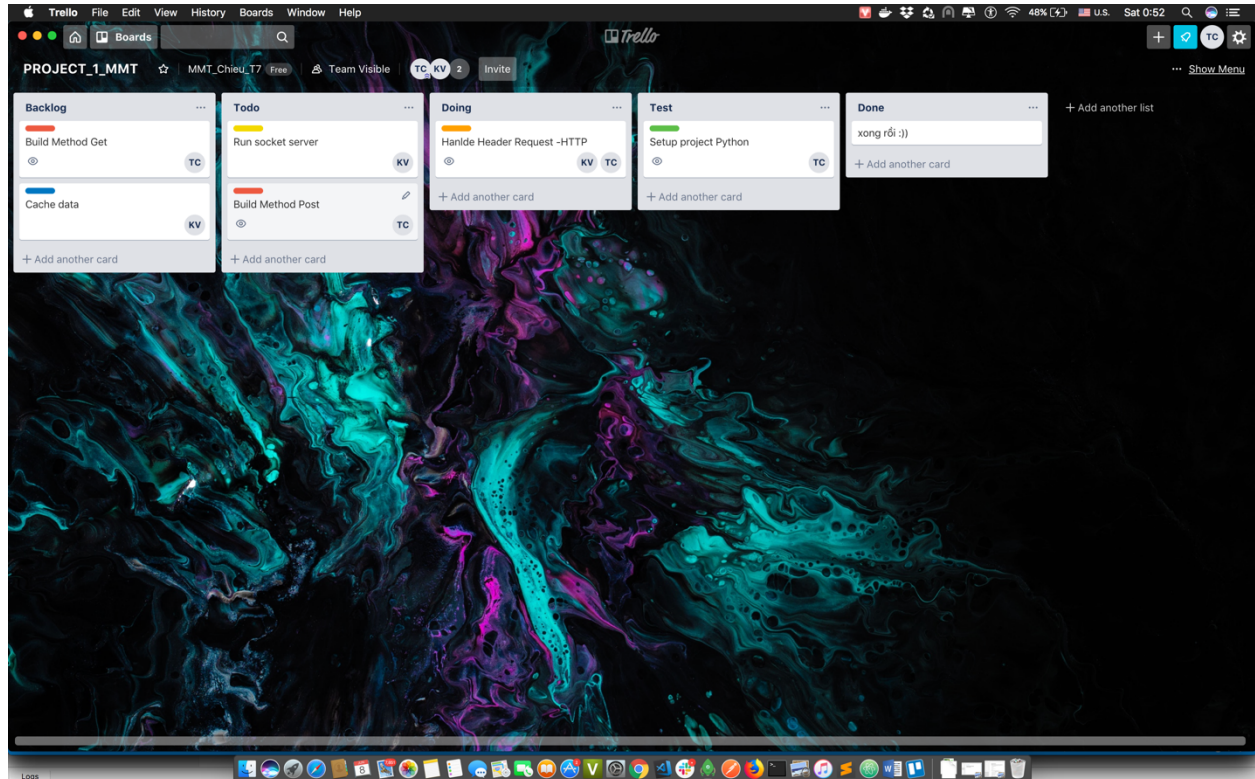
## Mục lục

I – THÀNH VIÊN VÀ PHÂN CÔNG CÔNG VIỆC .....	2
II – MÔ TẢ CÁCH THỨC CÀI ĐẶT HÀM .....	3
III – CHẠY KIỂM TRA CHƯƠNG TRÌNH VÀ KẾT QUẢ .....	6
IV – MÔ TẢ QUÁ TRÌNH NHẬN DỮ LIỆU GIỮ CLIENT – PROXY SERVER, PROXY SERVER – WEB SERVER.....	8
V - ỨNG DỤNG CỦA PROXY SERVER TRONG THỰC TẾ.....	10

## I – THÀNH VIÊN VÀ PHÂN CÔNG CÔNG VIỆC

HỌ TÊN	MSSV	CÔNG VIỆC	MỨC ĐỘ HOÀN THÀNH
Nguyễn Văn Tuấn	1512638	Quản lý dự án – Kiểm phát triển dự án	100 % công việc yêu cầu
Vũ Anh Linh	1712564	Phát triển dự án – Kiểm tester	100% công việc yêu cầu

Nhóm sử dụng công cụ quản lý dự án là Trello để quản lý dự án này. Việc sử dụng công cụ giúp nhóm cải thiện được hiệu quả làm việc và phân chia trách nhiệm rõ ràng. Dự án được quản lý theo mô hình agile-scrum chia nhỏ các task công việc và qui trình làm việc hiệu quả. Nhóm thường họp online để theo dõi tiến độ công việc của thành viên như thế nào, đặt mức deadline đúng để các tính năng không bị thiếu sót.



## II – MÔ TẢ CÁCH THỨC CÀI ĐẶT HÀM

STT	Class	Chức năng
1	<code>class DataParser(object)</code>	Xử lý dữ liệu để truyền đi
2	<code>class HttpParser(object):</code>	Convert header của http request
3	<code>class Connection(object):</code>	Tạo kết nối của client hoặc server để nhận và gửi dữ liệu đi
4	<code>class Server(Connection):</code>	Khởi tạo server
5	<code>class Client(Connection):</code>	Khởi tạo client
6	<code>class TCPServer(object):</code>	Khởi tạo socket giao thức tcp
7	<code>class ProxyServer(threading.Thread):</code>	Khởi tạo proxy server để xử lý các request
8	<code>class HTTP_PROTOCOL(TCP_SERVER):</code>	Xử lý request của giao thức http
9	<code>class Forbidden(Exception):</code>	Lỗi 403 Forbidden
10	<code>class ProxyError(Exception):</code>	Bắn lỗi proxy không connect được

STT	Hàm	Mô tả
1	<code>def run(self):TCP class</code>	<b>Chức năng:</b> Chạy server socket tcp port:8888 host: 127.0.0.1
2	<code>def handle_response(self, data)</code>	<b>Chức năng:</b> Gửi dữ liệu từ server về <b>Tham số:</b> data (Dữ liệu nhận từ server)
3	<code>def handle(self, client): HTTP_PROTOCOL class</code>	<b>Chức năng:</b> gửi request from client tới proxy threading <b>Tham số:</b> client (client server)
4	<code>def send(self, data)</code>	<b>Chức năng:</b> Gửi dữ liệu cho client hoặc server <b>Tham số:</b> data dữ liệu của client hoặc server
5	<code>def close(self): Connection class</code>	<b>Chức năng:</b> Đóng connect của client hoặc server
6	<code>def buffer_size(self)</code>	<b>Chức năng:</b> get size buffer của dữ liệu
7	<code>def flush(self)</code>	<b>Chức năng:</b> send dữ liệu đi gọi lại send()

8	<code>def processing(self)</code>	<b>Chức năng:</b> Trả dữ liệu cho client
9	<code>def parse(self, data)</code>	<b>Chức năng:</b> Convert header của client
10	<code>save_header_modified(details)</code>	<b>Chức năng:</b> lưu last modified của request nào đó phục vụ caching <b>Các tham số:</b> - Details: Dữ liệu client
11	<code>check_blocked(blocked, details)</code>	<b>Chức năng:</b> Kiểm tra domain có bị chặn không <b>Các tham số:</b> - Blocked: danh sách các domain bị block - Detail: Dữ liệu client
12	<code>res_status_forbiden(client_socket)</code>	<b>Chức năng:</b> Trả về màn hình lỗi 403 cho những domain bị block <b>Các tham số:</b> - Client_socket: client socket
13	<code>handle_request_from_client(client_socket, client_addr, client_data)</code>	<b>Chức năng:</b> xử lý request từ client (GET,POST) <b>Các tham số:</b> - Client_socket: client socket - Client_addr: địa chỉ client - Client_data: dữ liệu client
14	<code>proxy_start()</code>	<b>Chức năng:</b> khởi động proxy server để xử lý request từ client và trả về kết quả từ server-web
15	<code>method_get(client_socket, client_addr, details)</code>	<b>Chức năng:</b> Xử lý các GET request từ client - Client_socket: client socket - Client_addr: địa chỉ client - Client_data: dữ liệu client (method_post): tương tự
16	<code>do_cache_or_not(fileurl)</code>	<b>Chức năng:</b> kiểm tra có caching hay không <b>Các tham số:</b> - File_url: địa chỉ url để kiểm tra trong cache_data có file lưu chưa
17	<code>get_current_cache_info(fileurl)</code>	<b>Chức năng:</b> lấy thông tin địa chỉ file caching
18	<code>get_cache_details(client_addr, details)</code>	<b>Chức năng:</b> thông tin chi tiết của file caching để lưu vào header
19	<code>get_space_for_cache(fileurl)</code>	<b>Chức năng:</b> kiểm tra thông tin lưu lượng thư mục cache_data
20	<code>get_access(fileurl)</code>	<b>Chức năng:</b> phê duyệt để truy cập vào proxy server

21	<code>save_header_modified(details)</code>	<b>Chức năng:</b> lưu last modified của request nào đó phục vụ caching <b>Các tham số:</b> - Details: Dữ liệu client
22	<code>check_blocked(blocked, details)</code>	<b>Chức năng:</b> Kiểm tra domain có bị chặn không <b>Các tham số:</b> - Blocked: danh sách các domain bị block - Detail: Dữ liệu client
23	<code>res_status_forbidden(client_socket)</code>	<b>Chức năng:</b> Trả về màn hình lỗi 403 cho những domain bị block <b>Các tham số:</b> - Client_socket: client socket
24	<code>handle_request_from_client(client_socket, client_addr, client_data)</code>	<b>Chức năng:</b> xử lý request từ client (GET,POST) <b>Các tham số:</b> - Client_socket: client socket - Client_addr: địa chỉ client - Client_data: dữ liệu client
25	<code>method_get(client_socket, client_addr, details)</code>	<b>Chức năng:</b> Xử lý các GET request từ client - Client_socket: client socket - Client_addr: địa chỉ client - Client_data: dữ liệu client (method_post): tương tự
26	<code>do_cache_or_not(fileurl)</code>	<b>Chức năng:</b> kiểm tra có caching hay không <b>Các tham số:</b> - File_url: địa chỉ url để kiểm tra trong cache_data có file lưu chưa
27	<code>get_current_cache_info(fileurl)</code>	<b>Chức năng:</b> lấy thông tin địa chỉ file caching
28	<code>get_cache_details(client_addr, details)</code>	<b>Chức năng:</b> thông tin chi tiết của file caching để lưu vào header
29	<code>get_space_for_cache(fileurl)</code>	<b>Chức năng:</b> kiểm tra thông tin lưu lượng thư mục cache_data
30	<code>get_access(fileurl)</code>	<b>Chức năng:</b> phê duyệt để truy cập vào proxy server



### III – CHẠY KIỂM TRA CHƯƠNG TRÌNH VÀ KẾT QUẢ:

Danh sách test blacklist:

STT	Host
1	http://www.phimmoi.net
2	http://demo.unicore.asia/

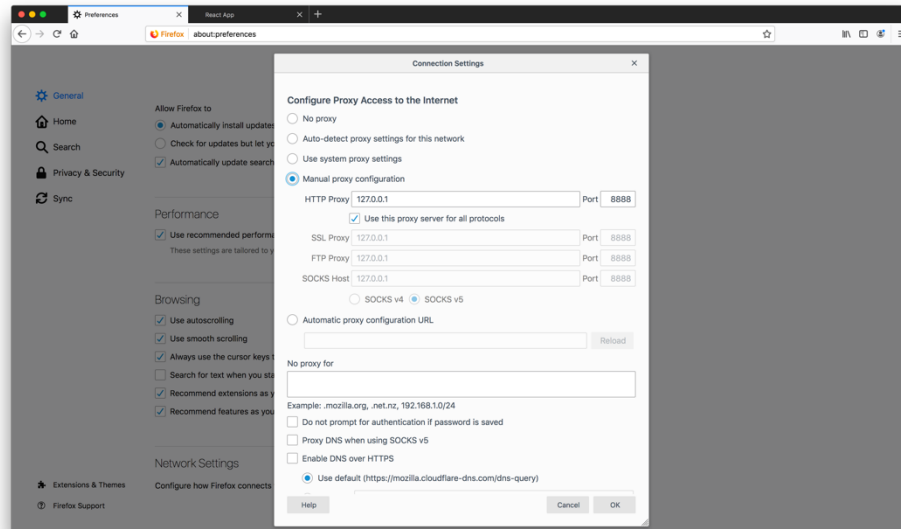
Hướng dẫn chạy file release: truy cập release folder rồi click vào file 1512638\_1712564.exe để chạy chương trình.

Các thức test chương trình sẽ được mô tả như sau:

**Bước 1: Setup proxy của firefox**

Host: 127.0.0.1 (localhost)

Port: 8888



Hình 1. Setup proxy trên firefox

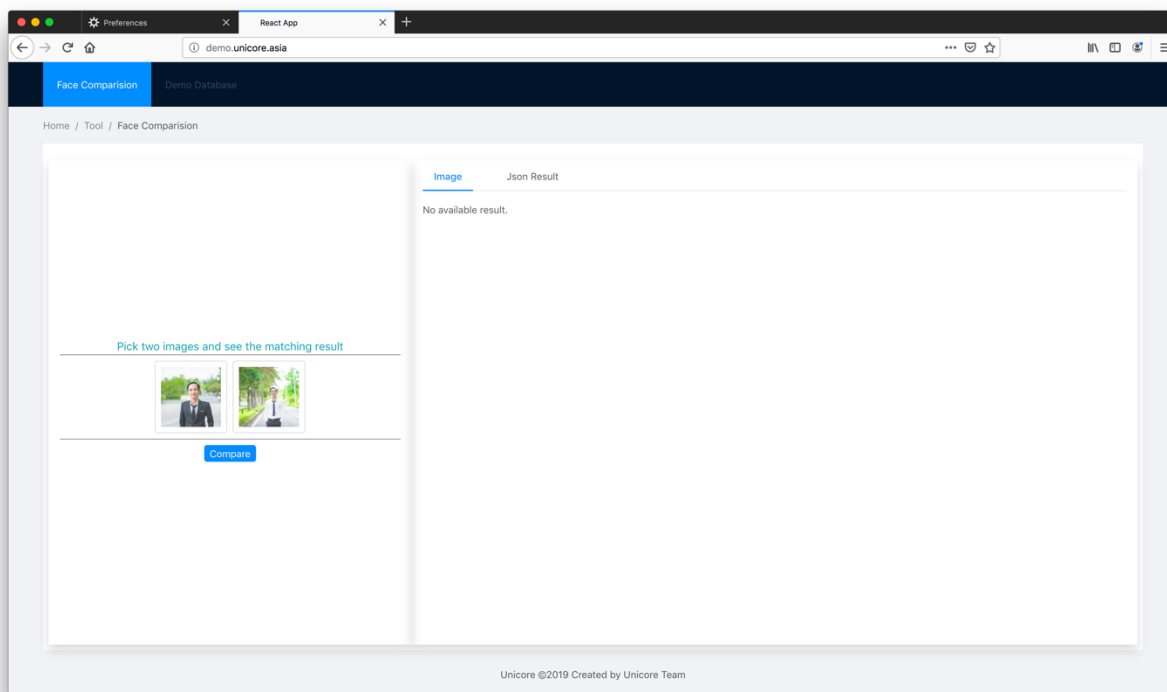
**Bước 2: Start server proxy**

```
Project_Last_1512638_1712564 — python3 Proxy_Server.py — python3 — Python Proxy_Server.py — 100x22
AC2019-06-12 16:06:11,224 - INFO - run:665 - Closing server socket
[→ Project_Last_1512638_1712564 python3 Proxy_Server.py
[*] Init sockets ... Finished
[*] Socket Binded successfully
[*] Server Started On Port Successfully [8888]

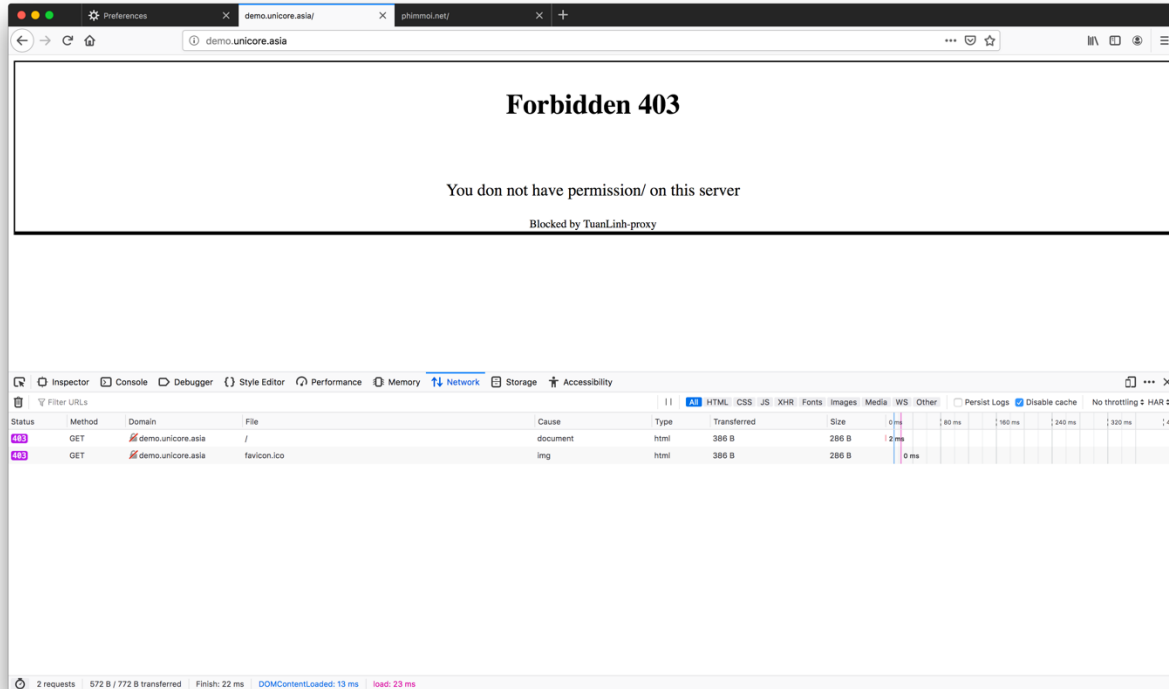
rcvd 1508 bytes from b'client'
rcvd 4344 bytes from b'server'
rcvd 8192 bytes from b'server'
rcvd 1025 bytes from b'server'
rcvd 0 bytes from b'server'
Closing proxy for connection <socket.socket [closed] fd=-1, family=AddressFamily.AF_INET, type=SocketKind.SOCK_STREAM, proto=0> at address ('127.0.0.1', 60120)
rcvd 219 bytes from b'client'
rcvd 1463 bytes from b'client'
rcvd 606 bytes from b'client'
rcvd 212 bytes from b'server'
rcvd 64 bytes from b'client'
rcvd 439 bytes from b'client'
rcvd 512 bytes from b'server'
rcvd 31 bytes from b'client'
rcvd 31 bytes from b'server'
```

Hình 2. Chạy proxy sever

### Bước 3: Truy cập web: và kết quả

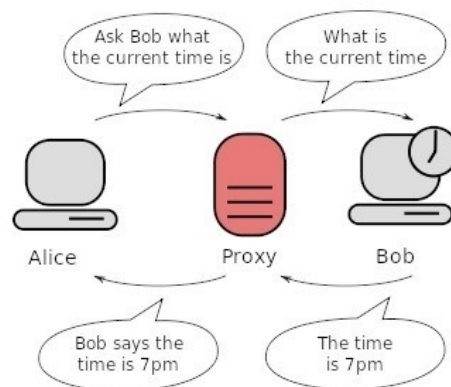


Hình 3. Truy cập một trang web có giao thức là http (<http://demo.unicore.asia/>)



Hình 4. Truy cập một trang web bị chặn – 403 Forbidden

#### IV – MÔ TẢ QUÁ TRÌNH NHẬN DỮ LIỆU GIỮA CLIENT – PROXY SERVER, PROXY SERVER – WEB SERVER



Trước tiên ta tìm hiểu proxy server là?

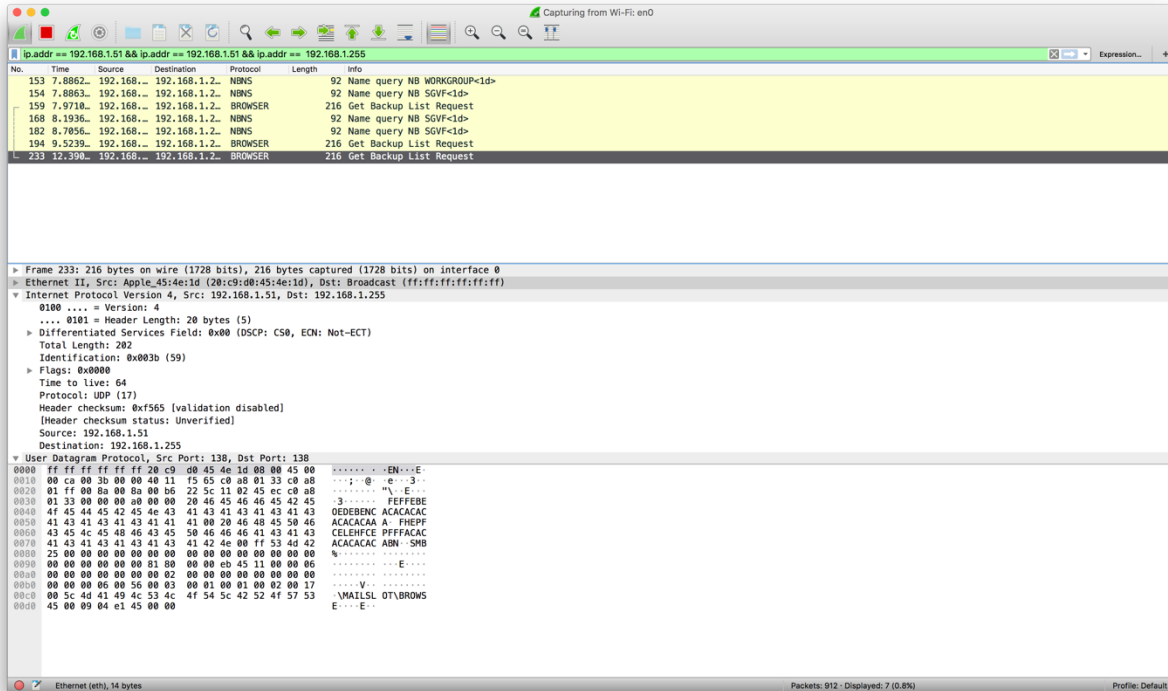
Proxy Server là một server đóng vai trò cài đặt proxy làm trung gian giữa người dùng trạm( workstation user) và Internet. Với Proxy Server, các máy khách( clients) tạo ra các kết nối đến các địa chỉ mạng một cách gián tiếp. Những chương trình client của người sử dụng sẽ qua trung gian proxy server thay thế cho server thật sự mà người sử dụng cần giao tiếp.

**Quá trình nhận dữ liệu giữa client-proxy server, giữa proxy server web server:**



## Đồ án Proxy Server Mạng Máy Tính

Trình duyệt (CLIENT) gửi GET http://SERVER/path HTTP/1.1 cho PROXY  
Bây giờ PROXY sẽ chuyển tiếp yêu cầu thực sự tới SERVER.  
Máy chủ sẽ chỉ xem PROXY là kết nối và trả lời cho PROXY giống như một CLIENT.  
PROXY nhận được phản hồi và chuyển nó trở lại CLIENT



## V. ỨNG DỤNG CỦA PROXY SERVER TRONG THỰC TẾ:

### *Ứng dụng trong thực tế:*

1. **Để kiểm soát việc sử dụng Internet của nhân viên và trẻ em:** Tổ chức và phụ huynh thiết lập máy chủ proxy để kiểm soát và giám sát nhân viên hoặc trẻ em sử dụng Internet. Ở các công ty bộ phận IT sẽ dùng proxy để chặn truy cập một số web như youtube hay Instagram, facebook để nhân viên không truy cập tập trung công việc hơn, đồng thời giám sát công việc của nhân viên (trừ lương...)
2. **Tiết kiệm băng thông và cải thiện tốc độ:** Các tổ chức cũng có thể nhận được hiệu suất mạng tổng thể tốt hơn khi sử dụng máy chủ proxy. Các máy chủ proxy có thể lưu vào bộ nhớ cache (lưu một bản sao trang web cục bộ) các trang web hay truy cập. Do đó khi yêu cầu trang tungtung.vn, máy chủ proxy sẽ kiểm tra xem có bản sao mới nhất của trang web này hay không và sau đó sẽ gửi cho bạn bản sao đã lưu. Điều này có nghĩa là khi hàng trăm người truy cập vào tungtung.vn cùng một thời điểm từ cùng một máy chủ proxy, máy chủ này chỉ cần gửi một yêu cầu đến tungtung.vn. Điều này giúp tiết kiệm băng thông của công ty và cải thiện hiệu suất mạng.
3. **Bảo mật riêng tư:** Cá nhân và tổ chức cũng sử dụng máy chủ proxy để duyệt Internet riêng tư hơn. Một số máy chủ proxy sẽ thay đổi địa chỉ IP và thông tin nhận dạng khác. Điều này có nghĩa là máy chủ đích không biết ai thực sự đã thực hiện yêu cầu ban đầu, giúp giữ thông tin cá nhân và thói quen duyệt web của bạn riêng tư hơn.
4. **Cải thiện bảo mật:** Bạn có thể cấu hình máy chủ proxy để mã hóa yêu cầu web để không ai có thể đọc được giao dịch của bạn. Ngoài ra, người dùng cũng có thể tránh các trang web độc hại thông qua máy chủ proxy. Các tổ chức có thể kết nối máy chủ proxy của họ với [Mạng riêng ảo \(VPN\)](#), do đó người dùng từ xa có thể truy cập Internet thông qua proxy của công ty. VPN kết nối trực tiếp đến mạng công ty để có thể kiểm soát và xác minh người dùng của họ có quyền truy cập vào các tài nguyên họ cần (email, dữ liệu nội bộ) đồng thời cũng cung cấp kết nối an toàn cho người dùng để bảo vệ dữ liệu công ty.(copy)
5. **Truy cập vào các tài nguyên bị chặn:** Máy chủ proxy cho phép người dùng phá vỡ các hạn chế nội dung do công ty hoặc một số tổ chức áp đặt. Nếu truy cập vào trang web bị chặn, bạn có thể đăng nhập vào máy chủ proxy ở nơi khác và xem từ đó. Máy chủ proxy khiến bạn giống như ở Mỹ nhưng thực ra bạn đang ở Việt Nam. Cái này phục vụ cho các nhu cầu riêng tư hoặc vì mục đích nào đó, bạn không thể truy cập một số web bị chặn bằng browser thông thường việc dùng proxy giúp bạn vượt qua dễ dàng

### *Lý do proxy server được sử dụng trong thực tế là:*

1. Do mọi thông tin truy xuất phải thông qua Proxy nên chúng ta có thể quản lý được mọi thông tin ra và vào ví dụ: Mọi yêu cầu của máy khách phải qua Proxy server, nếu địa chỉ IP có trên proxy, nghĩa là website này được lưu trữ cục bộ, trang này sẽ được truy cập mà không cần phải kết nối Internet, nếu không có trên Proxy server và trang này không bị cấm, yêu cầu sẽ được chuyển đến server thật, DNS server... và ra Internet.
2. Các dịch vụ proxy đều có lợi trong việc logging : Vì các proxy server hiểu các giao thức cơ bản, chúng cho phép logging đạt hiệu quả. Ví dụ, thay vì logging tất cả những dữ liệu đã truyền, một FTP (File Transfer Protocol) proxy server chỉ ghi lại những lệnh đã tạo và những đáp ứng của remote server, điều này giúp việc logging ít và hữu dụng hơn.
3. Đáp ứng được nhu cầu truy xuất của cá nhân và vừa đảm bảo an toàn cho hệ thống cục bộ do chúng ta sử dụng địa chỉ ẩn danh ,và mọi truy xuất đều thông qua proxy nên thông tin cục bộ không trực tiếp tương tác với bên ngoài.
4. Các dịch vụ proxy cho phép người dùng truy cập các dịch vụ Internet “trực tiếp”. Với các dịch vụ Proxy, các người dùng luôn nghĩ rằng họ đang tương tác trực tiếp với các dịch vụ Internet. Ví dụ các người dùng chỉ cần gõ vào địa chỉ của một trang web nào đó thì trang web được trình duyệt

hiển thị lên cho người dùng. Dĩ nhiên là có nhiều công việc phải làm ở bên trong nhưng nó là trong suốt đối với người dùng. Người dùng truy cập các dịch vụ Internet từ chính những hệ thống riêng của họ, mà không cần cho phép các gói tin truyền trực tiếp giữa hệ thống của người dùng và Internet đảm bảo an toàn cho hệ thống.

5. Proxy server tích lũy và cứu file , những file mà thường được yêu cầu bởi ngàn người dùng trên internet trong dữ liệu đặc biệt , gọi là cache . Do đó , proxy server chúng có thể tăng tốc độ truy nhập internet. Cache của proxy server có thể đã sẵn chứa thông tin bạn cần trong thời gian bạn yêu cầu , làm cho proxy server có thể phân phối thông tin ngay lập tức mà không cần phải truy tìm thông tin ngoài internet.
6. Một Proxy Server thường nằm bên trong tường lửa , giữa trình duyệt web và server thật , làm chức năng tạm giữ những yêu cầu Internet của các máy khách để chúng không giao tiếp trực tiếp Internet .Người dùng sẽ không truy cập được những trang web không cho phép ( bị công ty cấm ). Vd :Admin không muốn nhân viên của mình đọc báo hay chơi game online trong giờ làm việc , bằng cách dùng proxy server admin có thể khóa một số site được chỉ định.
7. Proxy server làm cho việc sử dụng băng thông có hiệu quả do chúng ta quản lý được các hoạt động của người dùng.Nên có thể giới hạn thông tin nào được dùng và không dùng tránh được việc nghẽn băng thông.

## V – TỔNG KẾT

Qua đồ án này nhóm hiểu hơn về hoạt động và chức năng của một server proxy là như thế, cách triển khai một proxy đơn giản để chạy các phương thức GET và POST. Cải thiện khả năng làm việc nhóm của các thành viên áp dụng những quy trình làm việc chuẩn mực để đạt hiệu suất công việc cao nhất

Đồ án có tham khảo một số nguồn trên github và các liên kết sau. Có những hàm được sử dụng lại và biến đổi phù hợp với nhu cầu của đồ án này. Để đảm bảo tính bản quyền những phần code được sử dụng thì sẽ có đánh dấu

Nguồn tham khảo:

<https://quantrimang.com/ly-thuyet-proxy-la-gi-117220>

<https://medium.com/@gdieu/build-a-tcp-proxy-in-python-part-1-3-7552cd5afdfc>

<https://stackoverflow.com/questions/41382127/how-does-the-python-socket-recv-method-know-that-the-end-of-the-message-has-been-received>

<https://null-byte.wonderhowto.com/how-to/sploit-make-proxy-server-python-0161232/>

<https://www.geeksforgeeks.org/creating-a-proxy-webserver-in-python-set-1/>

<https://www.geeksforgeeks.org/creating-a-proxy-webserver-in-python-set-2/>