

**TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN  
KHOA CÔNG NGHỆ THÔNG TIN**



**BÁO CÁO ĐỒ ÁN**

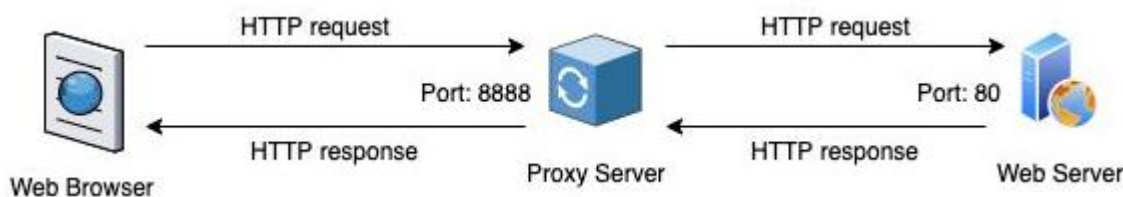
**ĐỀ TÀI**  
**LẬP TRÌNH SOCKET**  
**PROXY SERVER**

TÊN: Hoàng Anh Quân (19120628)  
: Mai Hồng Phúc (19120620)  
Trường ĐH KHTN

# LỜI MỞ ĐẦU

Trước hết cần hiểu Proxy là gì? Proxy Server (Máy chủ proxy) hoạt động như một cổng nối giữa người dùng và Internet. Đây là một server trung gian giữa người dùng cuối và trang web họ truy cập. Các máy chủ proxy cung cấp các chức năng, bảo mật và riêng tư khác nhau phụ thuộc vào nhu cầu của bạn hoặc chính sách công ty.

Nguyên tắc hoạt động cơ bản của proxy Server là : Proxy server xác định những yêu cầu từ phía client và quyết định đáp ứng hay không đáp ứng, nếu yêu cầu được đáp ứng, proxy server sẽ kết nối tới server thật thay cho client và tiếp tục chuyển tiếp đến những yêu cầu từ client đến server, cũng như đáp ứng những yêu cầu của server đến client.



Mô hình proxy

Dựa theo mô hình trên chúng tôi đã tiến hành phát triển chương trình ProxyServer viết bằng ngôn ngữ python3 với giao diện console kinh điển.

Chương trình thực hiện được các chức năng cơ bản của một máy chủ proxy như lọc các gói tin, chặn các trang web cấm, phân tích gói tin, xử lý đa luồng,... thông qua giao thức HTTP.

Soure code của máy chủ Proxy Server gồm 3 file: ProxyServer.py chứa nội dung chương trình; 403.html chứa nội dung của mã lỗi 403 Forbidden; blacklist.conf chứa thông tin của các host bị cấm.

Dưới đây là nội dung phần báo cáo trình bày những nội dung làm được và cách để chạy chương trình Proxy cũng như đánh giá mức độ hoàn thành của chương trình. Chúng tôi sẽ cố gắng phát triển để chương trình ngày càng hoàn thiện hơn. Xin cảm ơn.

# NỘI DUNG BÁO CÁO

## I. PHÂN CÔNG CÔNG VIỆC

### 1. MSSV: 19120620

Mô tả công việc	<ul style="list-style-type: none"><li>- Viết các hàm tạo các socket kết nối từ ProxyServer đến client và web server, các hàm đọc request và response, hàm in request.</li><li>- Giải thích lí do cần Proxy Server trong thực tế.</li><li>- Kiểm thử chương trình, và phát hiện lỗi và sửa chữa lỗi.</li><li>- Tổng hợp, viết báo cáo.</li></ul>
-----------------	---

### 2. MSSV: 19120628

Mô tả công việc	<ul style="list-style-type: none"><li>- Viết các hàm xử lí request, trả về mã lỗi, kiểm tra hostname trong blacklist, hàm ProcessProxy và hàm main xử lí đa luồng.</li><li>- Chạy chương trình ProxyServer và chụp hình, bắt gói tin wireshark và phân tích quá trình truyền nhận dữ liệu giữa Client - Proxy Server và Proxy Server – Web Server.</li><li>- Chỉnh sửa và hoàn thành báo cáo.</li></ul>
-----------------	---

## II. CÁC HÀM CHỨC NĂNG TRONG CHƯƠNG TRÌNH

(Ngôn ngữ lập trình: **Python3**)

### 1. CreateServer(host:str, port:int)

- Input: host: hostname (string)  
port: một số nguyên (int)
- Output: Hàm trả về đối tượng socket gắn với địa chỉ host và cổng port, socket này dùng để kết nối với client.
- Chức năng: Tạo một socket server có địa chỉ host và cổng port.

### 2. CreateClient(host:str, port:int)

- Input: host: hostname (string)  
port: một số nguyên (int)



- Output: Hàm trả về đối tượng socket kết nối từ proxy server tới web server có địa chỉ `host` và cổng `port`.
  - Chức năng: Tạo một socket client kết nối tới địa chỉ `host` và cổng `port`.
- 3. `parseRequest(ReqMsg:str)`**
- Input: `ReqMsg`: yêu cầu nhận từ client (str)
  - Output: Hàm trả về 1 dictionary chứa các key `method`, `protocol`, `host`, `path`, `version` đã được phân tích từ yêu cầu đã nhận.
  - Chức năng: Phân tích dòng đầu tiên của request
- 4. `printReqInfo(Inf:dict)`**
- Input: `Inf`: dictionary chứa các key đã được phân tích từ yêu cầu của web browser.
  - Output: Hàm in ra các key trong dictionary.
  - Chức năng: In thông tin cơ bản của Request lên màn hình console
- 5. `Send403Code(sockfd:socket)`**
- Input: `sockfd`: là 1 đối tượng socket kết nối tới client
  - Output: Gửi mã lỗi 403 về cho client
  - Chức năng: Gửi data chứa mã lỗi 403 cho web browser, browser sẽ hiện thị mã lỗi đã lên màn hình
- 6. `checkBlacklist(host:str,filename:str)`**
- Input: `host`: hostname của website cần kiểm tra (string)  
`filename`: tên của file blacklist (string)
  - Output: Trả về True nếu `host` tồn tại trong blacklist, trả về False nếu ngược lại
  - Chức năng: Hàm kiểm tra `host` có tồn tại trong blacklist hay không
- 7. `readRequest(sock:socket)`**
- Input: `sock`: là một đối tượng socket kết nối với client
  - Output: Hàm trả về yêu cầu của client gửi tới Proxy Server.
  - Chức năng: Nhận các gói dữ liệu từ browser gửi lên cho máy chủ proxy
- 8. `readResponse(sock:socket)`**
- Input: `sock`: là một đối tượng socket kết nối với web server
  - Output: Hàm trả về hồi đáp của web server gửi tới Proxy Server.
  - Chức năng: Nhận các gói dữ liệu từ webServer gửi lên cho máy chủ proxy
- 9. `ProcessProxy(proxyServer:socket)`**
- Input: `proxyServer`: là một đối tượng socket kết nối với client

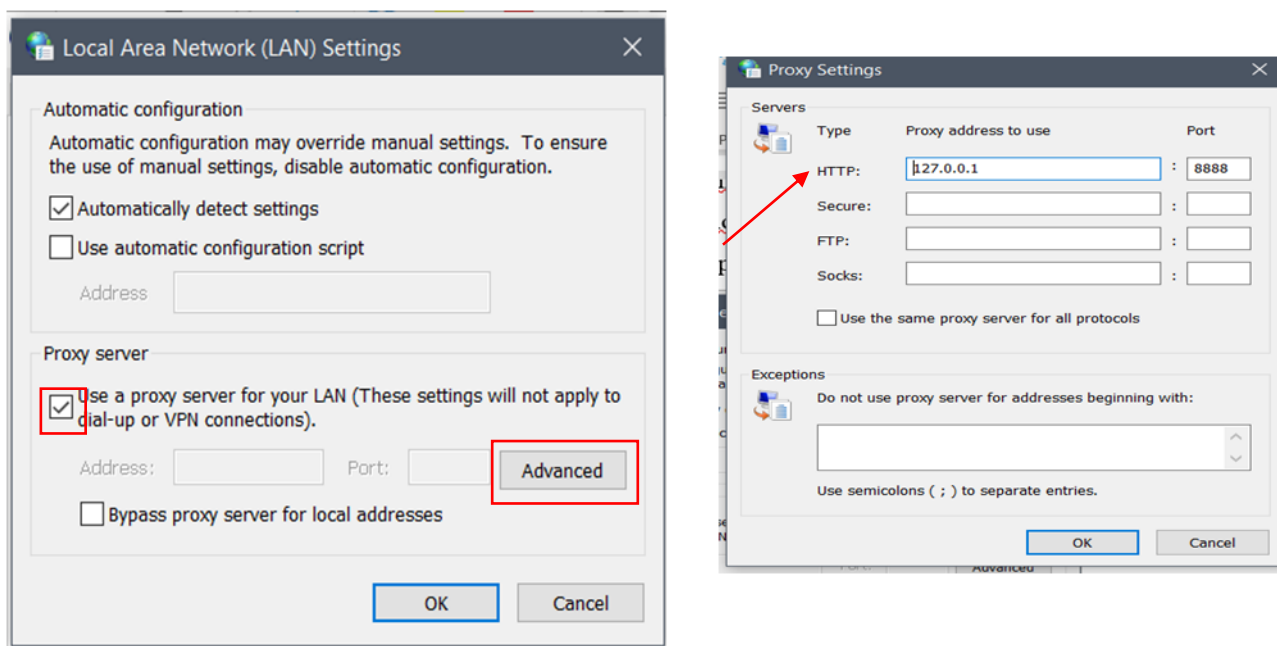
- Output: Trả về 1 nếu xử lý thành công và trả về không nếu có lỗi.
- Chức năng: Là hàm làm nhiệm vụ xử lý chính cho mọi sự kiện nhận gửi các dữ liệu từ web browser ->Proxy Server ->Webserver và ngược lại, cũng như xử lý các gói dữ liệu đầu vào và đầu ra và chặn các host nằm trong Blacklist.

### III. QUÁ TRÌNH CHẠY PROXY

Trước khi thực hiện chạy chương trình, đầu tiên bạn phải cài đặt python3 từ trang chủ python. Sau đó thực hiện các bước sau:

#### 1. Cấu hình Proxy với protocol HTTP

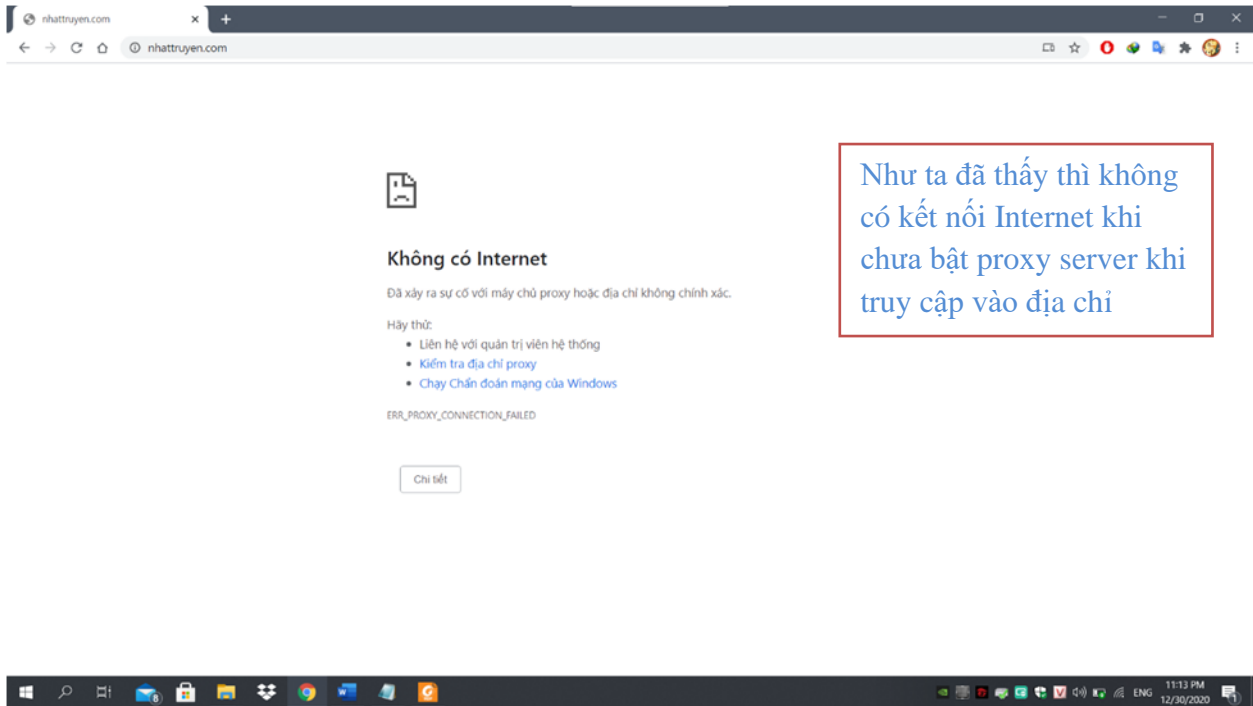
- Mở hộp thoại **Local Area Network (LAN) Settings** thông qua đường dẫn: **Internet Options → Connections → LAN settings**.
- Sau đó tích vào trường **Use a proxy server for your LAN** và click vào **Advance**.
- Ngay sau đó hộp thoại **Proxy Setting** được hiện ra. Nhập địa chỉ IP cho Proxy Sever và số hiệu Port cho trường **HTTP** và nhấn **OK** để kết thúc việc cấu hình.



Hình 1: Cấu hình cho proxy đến proxy server

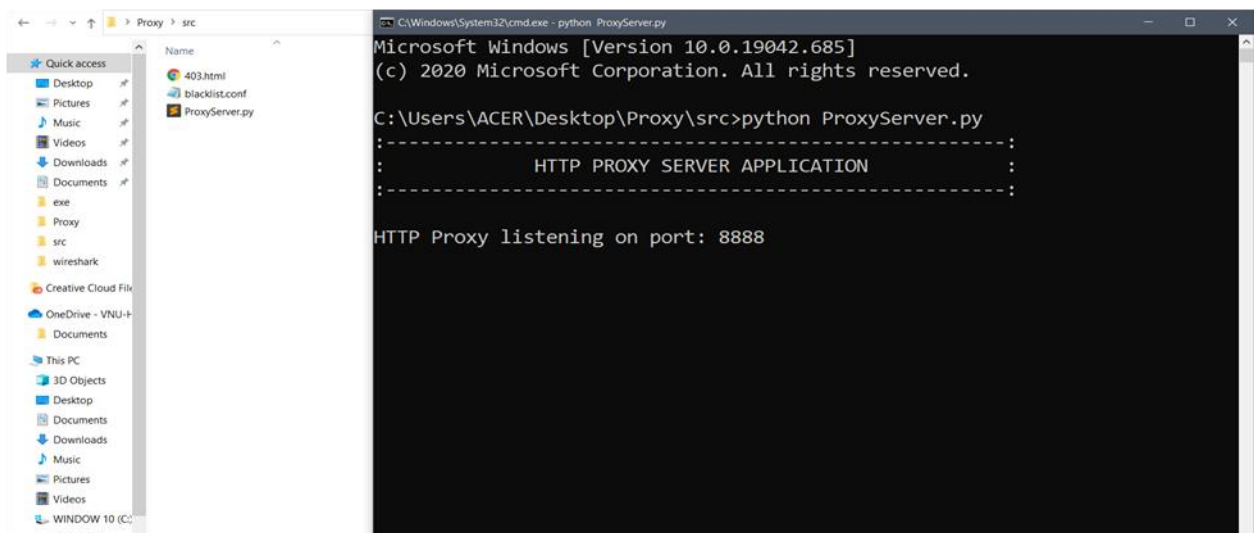
## 2. Bật máy chủ Proxy và xem kết quả.

- Trước khi bật máy chủ proxy chúng ta thử kiểm tra xem có truy cập vào được một website **HTTP** nào đó không (ví dụ: <http://nhattruyen.com>).



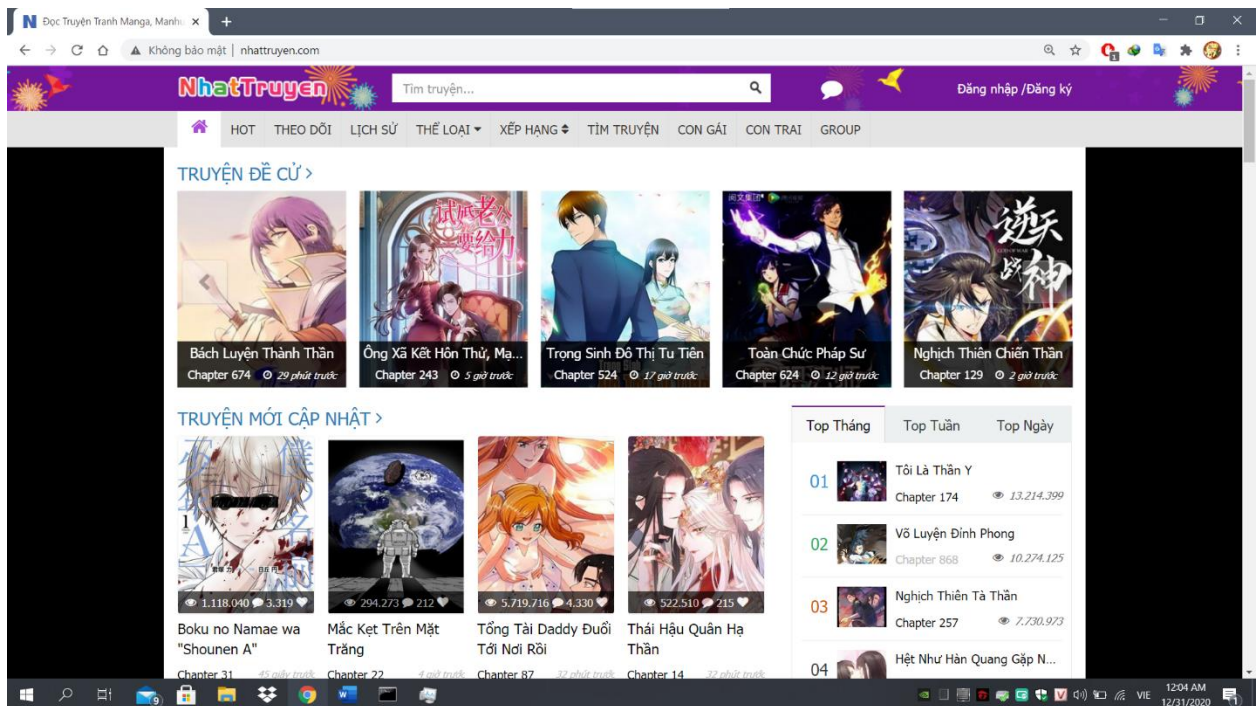
Hình 2: Kiểm tra kết nối khi chưa bật Proxy Sever

- Mở proxy server thông qua dòng lệnh **python ProxyServer.py** vào đường dẫn chứa thư tệp **ProxyServer.py** trên **CMP**. Giao diện máy chủ proxy hiện ra như hình 3.



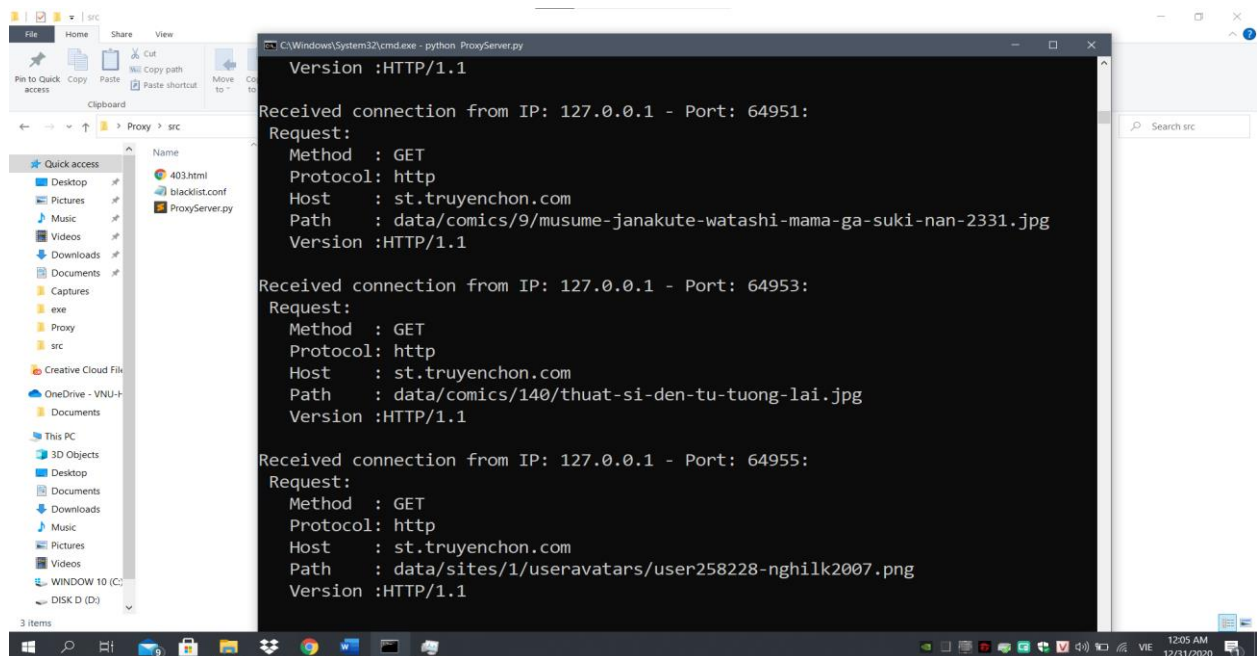
Hình 3: Mở Proxy Server

- Thử truy cập lại website <http://nhattruyen.com>.  
Kết quả chạy được trên web browser:



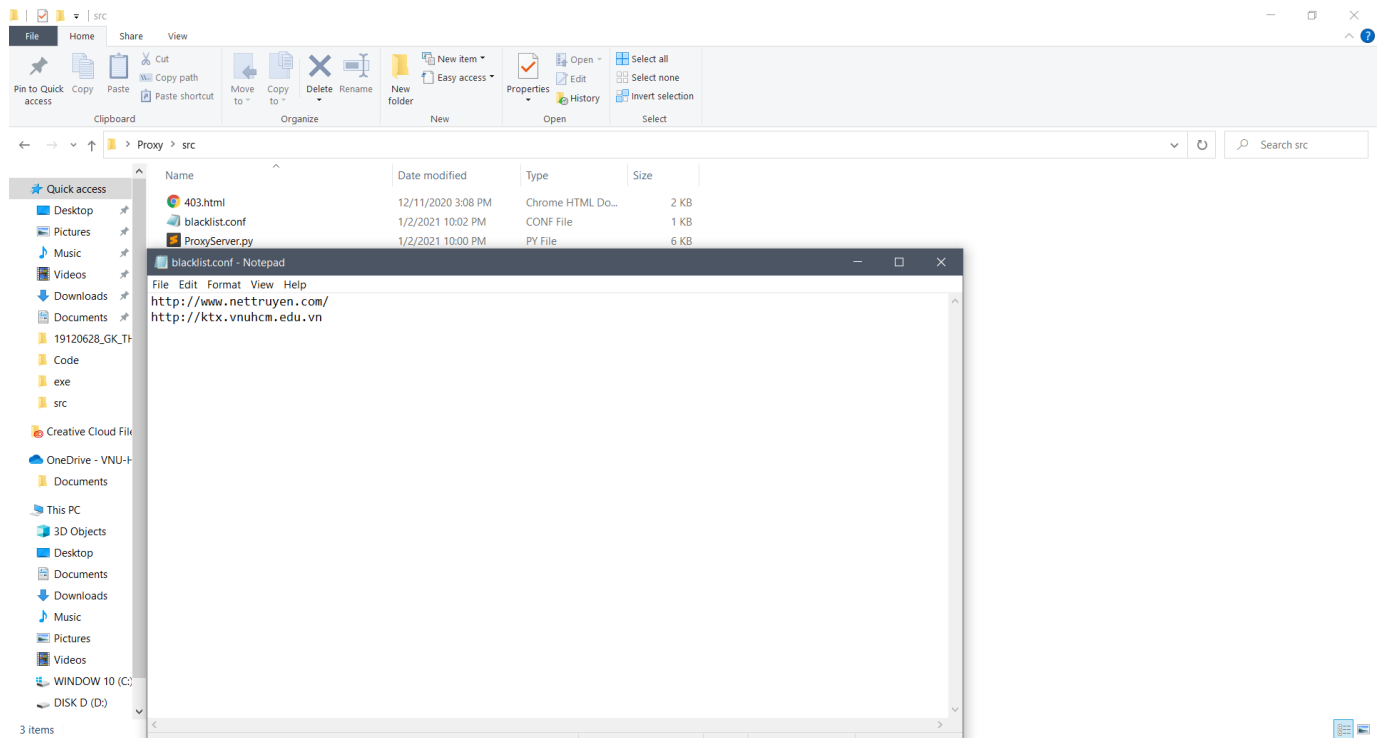
Hình 4: Truy cập vào website nhattruyen.com sau khi bật proxy

Kết quả chạy được trên Proxy Server:



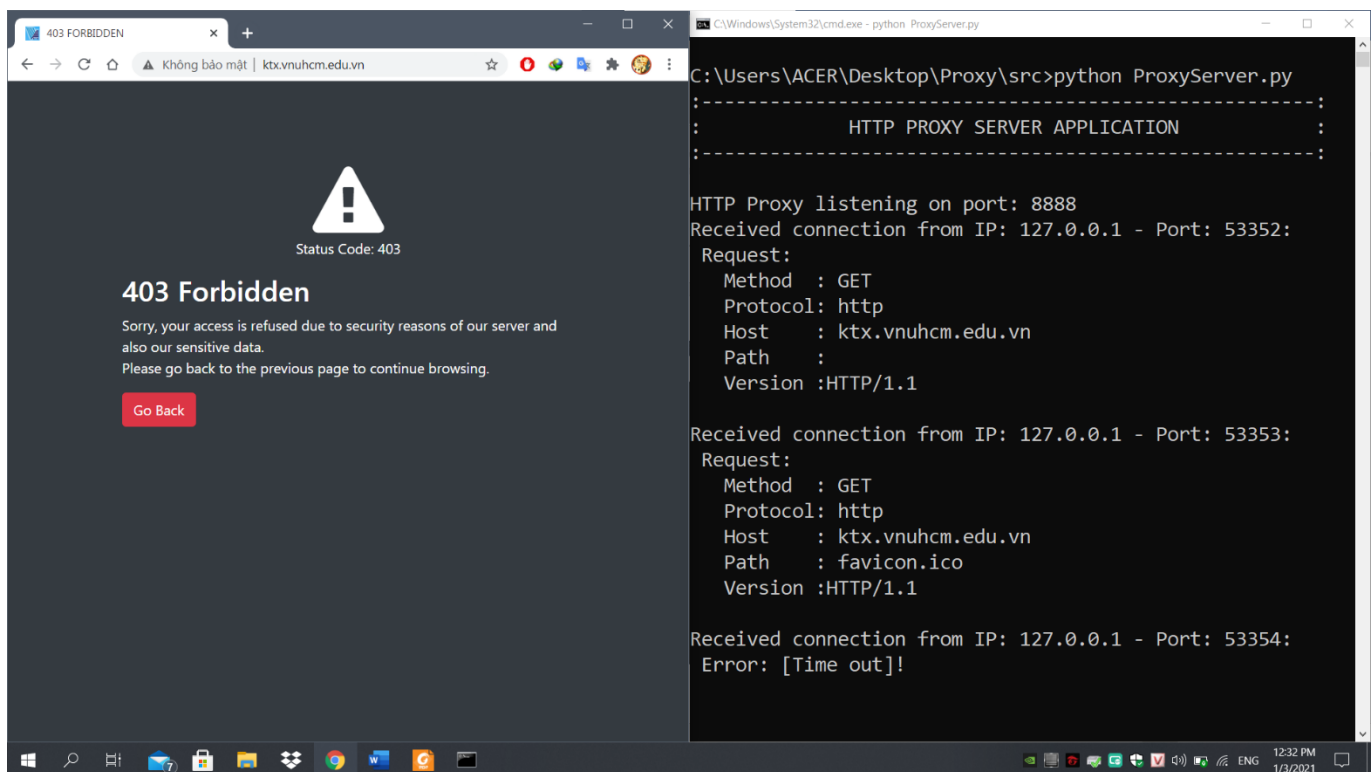
Hình 5: Dữ liệu nhận vào được in ra trên máy chủ Proxy.

- Chạy một website nằm trong **blacklist.conf** (ví dụ: **ktx.vnuhcm.edu.vn**)



Hình 6: blacklist

Kết quả nhận được: Mã lỗi **403 Forbidden**



Hình 7: Truy cập vào trang web nằm trong blacklist



## IV. MỨC ĐỘ HOÀN THÀNH

Những công việc mà chương trình ProxyServer có thể làm được:

1. Proxy Server cho phép Client truy cập website thông qua các method: GET, POST.
2. Proxy Server hỗ trợ HTTP, không hỗ trợ HTTPS.
3. Proxy Server xử lý đồng thời được các request từ client.
4. Proxy Server lắng nghe tại port 8888, chờ kết nối từ client.
5. Proxy Server chặn được tất cả các truy cập đến các domain website có trong file **blacklist.conf** và trả về status code 403 Forbidden về cho client.

**Tự đánh giá mức độ hoàn thành: 100%**

## V. PHÂN TÍCH CÁC GÓI TIN BẮT ĐƯỢC QUA WIRESHARK

## VI. LÝ DO CẦN PROXY TRONG THỰC TẾ

### 1. Cải thiện an ninh:

Nó cho phép ẩn địa chỉ IP gốc của hệ thống máy khách, giúp máy khách lướt internet với một địa chỉ IP ẩn danh. Máy chủ proxy đảm bảo rằng các chi tiết của khách hàng được an toàn trước các cuộc tấn công và truy cập không mong muốn.

### 2. Kiểm soát truy cập Internet:

Nó sử dụng máy chủ Proxy để thực hiện kiểm soát truy cập Internet như xác thực cho kết nối Internet, kiểm soát thời gian trực tuyến, lọc web Internet, kiểm soát băng thông và bộ lọc nội dung, v.v.

### 3. Bỏ qua các hạn chế và bộ lọc bảo mật:

Proxy giúp khách hàng sử dụng bỏ qua các hạn chế truy cập internet và truy cập các trang web bị chặn dễ dàng.

## MỤC LỤC

I. PHÂN CÔNG CÔNG VIỆC.....	3
1. MSSV: 19120620.....	3
2. MSSV: 19120628.....	3
II. CÁC HÀM CHỨC NĂNG TRONG CHƯƠNG TRÌNH .....	3
1. CreateServer(host:str, port:int) .....	3
2. CreateClient(host:str, post:int).....	3
3. parseRequest(ReqMsg:str).....	4
4. printReqInfo(inf:dict).....	4
5. Send403Code(sockfd:socket) .....	4
6. checkBlacklist(host:str,filename:str) .....	4
7. readRequest(sock:socket) .....	4
8. readResponse(sock:socket).....	4
9. ProcessProxy(proxyServer:socket).....	4
III. QUÁ TRÌNH CHẠY PROXY .....	5
1. Cấu hình Proxy với protocol HTTP.....	5
2. Bật máy chủ Proxy và xem kết quả. ....	6
IV. MỨC ĐỘ HOÀN THÀNH.....	9
V. PHÂN TÍCH CÁC GÓI TIN BẮT ĐƯỢC QUA WIRESHARK.....	9
VI. LÍ DO CẦN PROXY TRONG THỰC TẾ .....	10
1. Cải thiện an ninh: .....	10
2. Kiểm soát truy cập Internet:.....	10
3. Bỏ qua các hạn chế và bộ lọc bảo mật:.....	10

THE END