

TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN
KHOA: Công nghệ thông tin

Báo cáo đồ án

Socket

Mạng máy tính | 2019



GVHD: Lê Hà Minh

Sinh viên:

1712419 – Nguyễn Hữu Hào
1712837 – Mạng Viên Triệu
1712827 – Trần Quốc Toàn
1712844 – Hồ Nguyễn Chí Trung

MỤC LỤC

1. Thông tin sinh viên.....	3
2. Phân công công việc.....	3
3. Các hàm chức năng.....	5
a) Hàm <code>wchar_t* Convert(const char* charArray)</code>	5
b) Hàm <code>void Star_load_blackList(vector<string> &array)</code>	5
c) Hàm <code>bool Check_sever_name_inBlackList(string severname)</code>	5
d) Hàm <code>void outputfile(char *filename, string &a)</code>	5
e) Hàm <code>void Get_port(string &buf, string &Address, int &Port)</code>	5
f) Hàm <code>unsigned int Get_down(void *m)</code>	5
g) Hàm <code>void str_Split(string str, vector<string> &conten, char deim)</code>	6
h) Hàm <code>sockaddr_in*Get_IP_Sefer(stringTenserver,char*Domain_Name)</code>	6
k) Hàm <code>void CloseSocket()</code>	6
i) Hàm <code>unsigned int Proxy_to_websever(void*m)</code>	6
j) Hàm <code>unsigned int Client_to_proxy(void*m)</code>	6
m) Hàm <code>void StartSocket()</code>	6
4. Chạy chương trình và kết quả.....	6
5. Đánh giá mức độ hoàn thành.....	9
6. Bắt gói tin và quá trình nhận gửi dữ liệu của proxy-sever.....	10
7. Trong thực tế, tại sao cần proxy?.....	12
8. Ngôn ngữ lập trình sử dụng và công cụ hỗ trợ.....	14
9. Tài liệu tham khảo.....	15

1. Thông tin sinh viên

STT	MSSV	Họ Tên	Email
1	1712419	Nguyễn Hữu Hào	1712419@student.hcmus.edu.vn
2	1712837	Mang Viên Triệu	1712837@student.hcmus.edu.vn
3	1712844	Hồ Nguyễn Chí Trung	1712844@student.hcmus.edu.vn
4	1712827	Trần Quốc Toàn	1712827@student.hcmus.edu.vn

2. Phân công công việc

a) Yêu cầu đề bài:

1. Hỗ trợ HTTP 1.0 và HTTP 1.1
2. Chương trình cho phép Client truy cập website thông qua Proxy Server (GET, POST)
3. Chương trình chỉ cần hỗ trợ http, không cần https.
4. Proxy Server phải xử lý đồng thời được các request từ client.
5. Proxy Server sẽ chạy trên port 8888.
6. Proxy Server cấu hình 1 file blacklist.conf, mỗi dòng chứa các domain website cấm Client truy cập. Proxy Server phải chặn tất cả các truy cập trùng với các domain trong file.
Ví dụ: abc.com.vn
www.thongtin.net ...
7. Proxy Server có chức năng caching lại những nội dung của trang web, để phục vụ cho Client khác nhanh hơn. Sinh viên tự đưa ra cơ chế caching.

Yêu cầu thực hiện của nhóm:

- Đảm bảo đáp ứng đầy đủ các yêu cầu trong mô tả đồ án.
- Đảm bảo đúng deadline đã giao.
- Có kế hoạch kiểm tra tiến độ thường xuyên.
- Có hình ảnh minh họa.

b) Phân công công việc

MSSV	Họ Tên	Công việc	Đánh giá
1712419	Nguyễn Hữu Hào	<ul style="list-style-type: none">- Xử lý đa luồng.- Tạo danh sách các domain cần.- Kiểm tra domain với các domain trong blacklist.- Viết html in ra 403 forbidden	100%
1712837	Mang Viên Triệu	<ul style="list-style-type: none">- Viết báo cáo.- Viết hàm tách host name.- Chuyển http 1.1 sang 1.0	100%
1712827	Trần Quốc Toàn	<ul style="list-style-type: none">- Xử lý đa luồng.- Viết hàm lấy address và port.- Lấy IP từ domain.	100%
1712844	Hồ Nguyễn Chí Trung	<ul style="list-style-type: none">- Tìm hiểu tại sao cần proxy-sever- Xử lý đa luồng.- Tổng hợp	100%

3. Các hàm chức năng

a) **Hàm** `wchar_t* Convert(const char* charArray)`

- Chức năng: Chuyển `char*` sang `wchar*`
- Kết quả: Trả về giá trị `wchar*`

b) **Hàm** `void Star_load_blackList(vector<string> &array);`

- Hàm chạy file `Blacklist.inf`. Nếu không tồn tại thì cout ra “Blacklist cannot load!!!”.
- Yêu cầu phải để file `Blacklist` chung với `proxy-sever.exe`.

c) **Hàm** `bool Check_sever_name_inBlackList(string severname);`

- Hàm này sẽ tiến hành kiểm tra những địa chỉ mà ta cần tìm trong nằm trong blacklist hay.
- Hàm trả về giá trị 1 hoặc 0.

d) **Hàm** `void outputfile(char *filename, string &a);`

- Hàm này tiến học đọc file `forbidden403` (code html) lên rồi tiến hành lưu vào chuỗi `string Prohibit_access`(chuỗi in lên Thông báo 403 forbidden).
- Có thể sử dụng hàm này để lưu vào string a hoặc nhập tay trực tiếp vào string a.
- Hàm trả về chuỗi html được lưu trong a

e) **Hàm** `void Get_port(string &buf, string &Address, int &Port);`

- Hàm Lấy Address và port đồng thời tạo truy vấn từ truy vấn nhận được từ Client truyền tới.
- Truyền vào chuỗi buffer qua hàm này thì sẽ lấy được địa chỉ domain và port 80.

f) **Hàm** `unsigned int Get_down(void *m);`

Hàm Nhập Input từ màn hình Console Nếu muốn dừng lại thì nhập end vào console.

g) Hàm `void str_Split(string str, vector<string> &conten, char deim);`

Hàm tách chuỗi str truyền vào. Sau đó tiến hành push_back vào conten.

h) Hàm `sockaddr_in* Get_IP_Seiver(string Tenserver, char*Domain_Name);`

- Hàm tiến lấy địa chỉ IP để yêu cầu kết nối.

- Hàm trả về một sockaddr_in.

k) Hàm `void CloseSocket();`

Hàm Đóng các sever lại (đóng cổng giao tiếp lại).

i) Hàm `unsigned int Proxy_to_websever(void *m);`

Hàm Thread giao tiếp giữa Proxy Server và WebServer(Remote Server).

j) Hàm `unsigned int Client_to_proxy(void *m);`

Hàm thread Giao tiếp của client đến proxy thực hiện truy cập từ client đến proxy.

m) Hàm `void StartSocket();`

Hàm Bắt đầu khởi tạo Server để lắng nghe các kết nối từ client chuyển đến cho proxy.

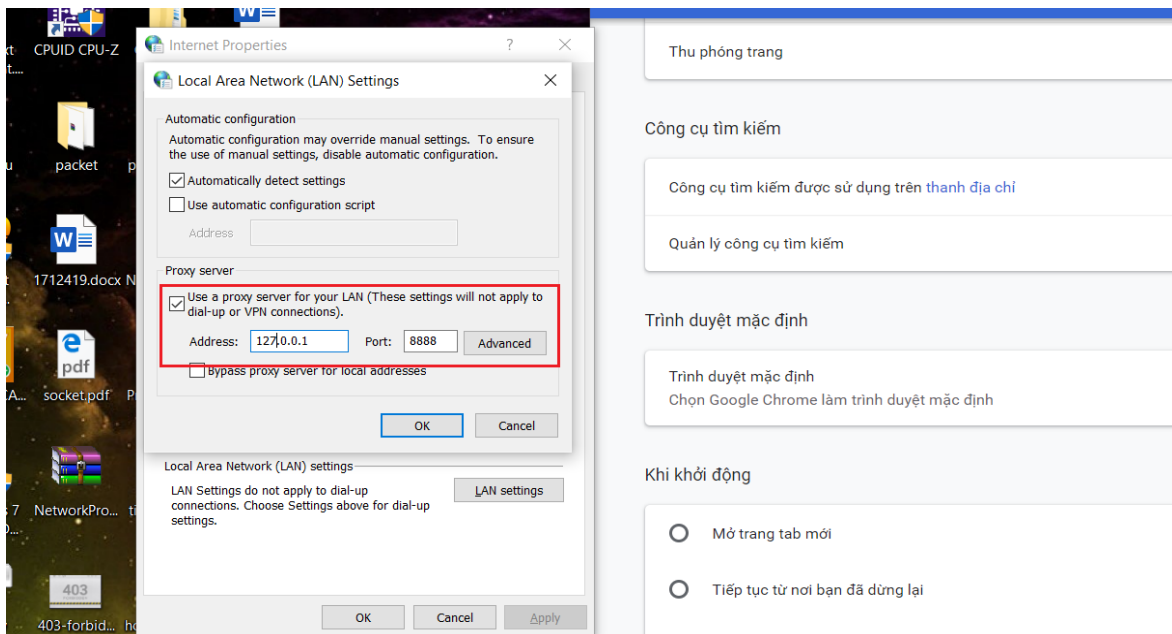
4. Chạy chương trình và kết quả.

Đầu tiên ta tiến hành cài đặt proxy của máy (Em sử dụng google chrome).

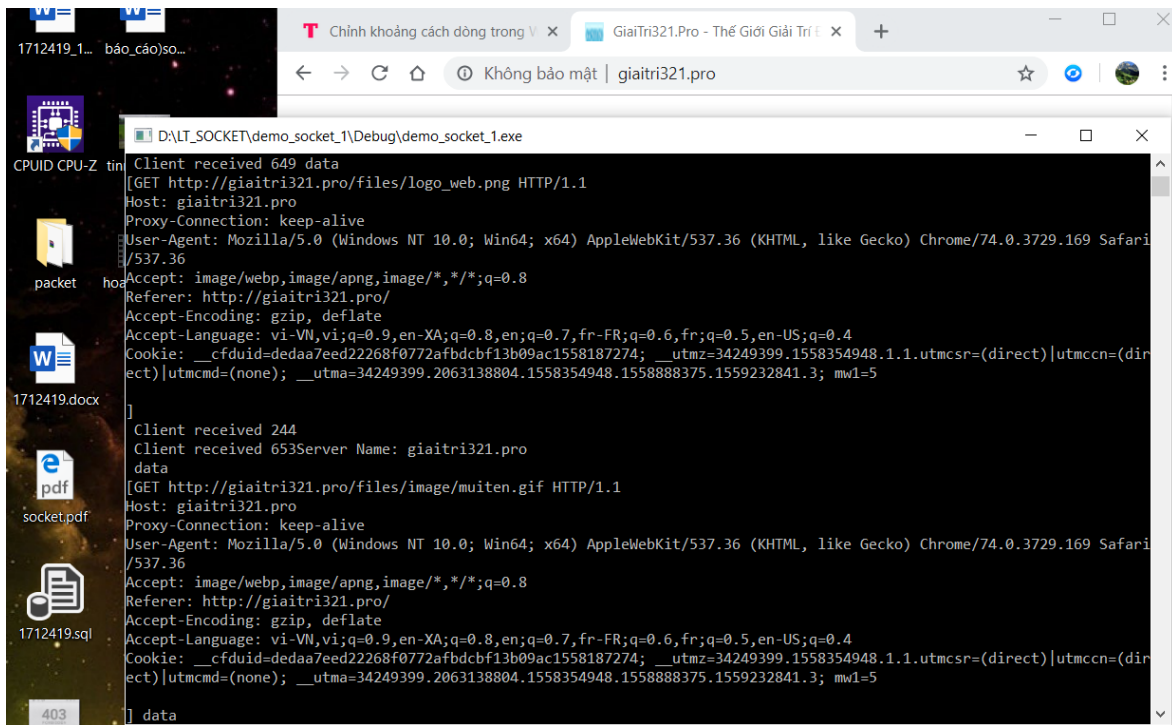
- Address: 127.0.0.1

- Port: 8888

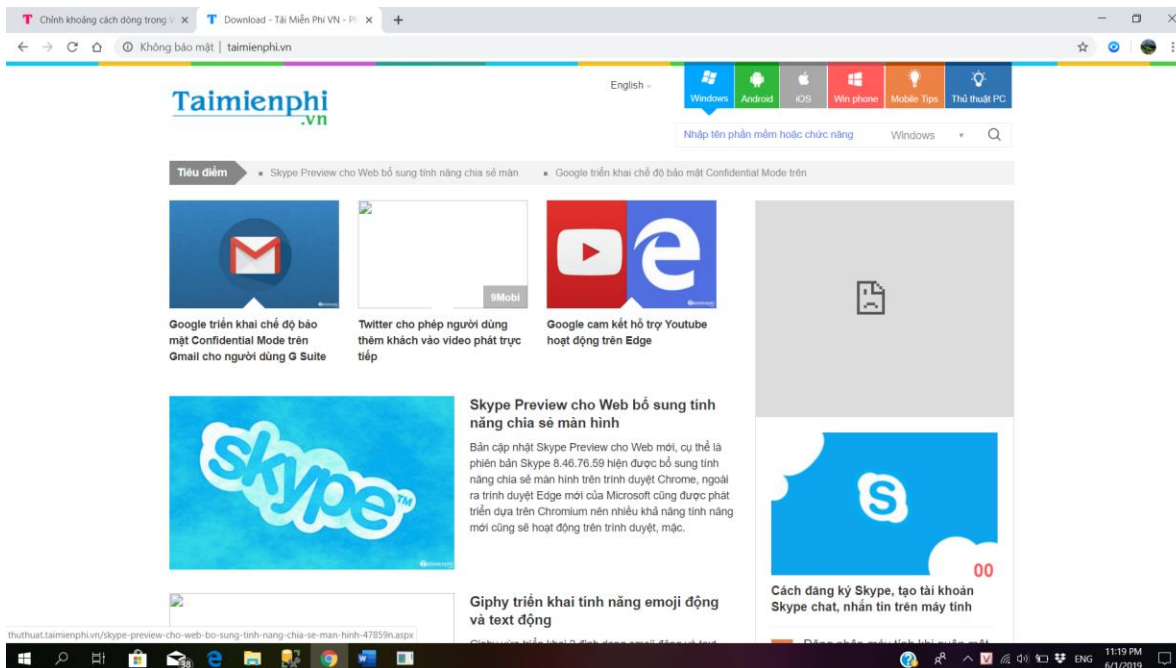
Mở trình duyệt lên, sau đó vào cài đặt, ấn chữ proxy vào phần tìm kiếm, sau đó chọn vào Mở cài đặt proxy.



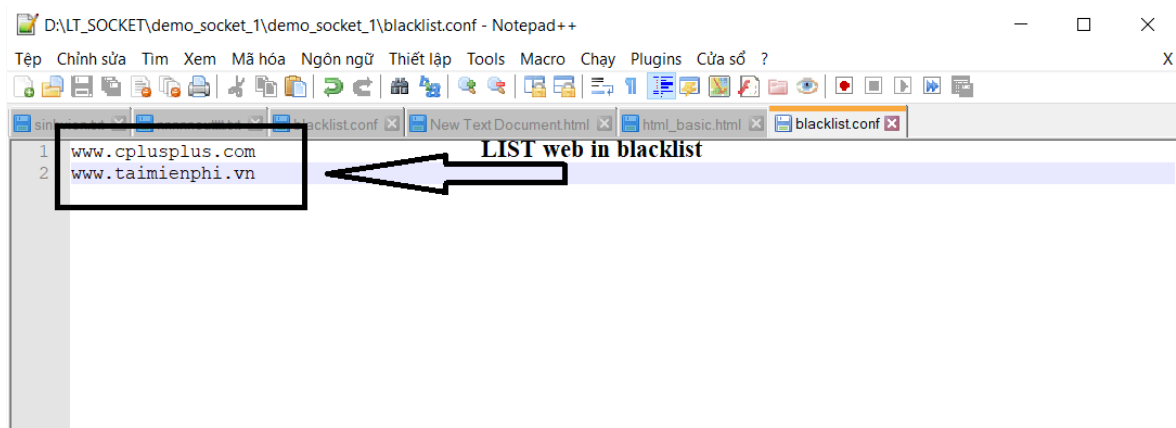
Các gói tin Server nhận được:



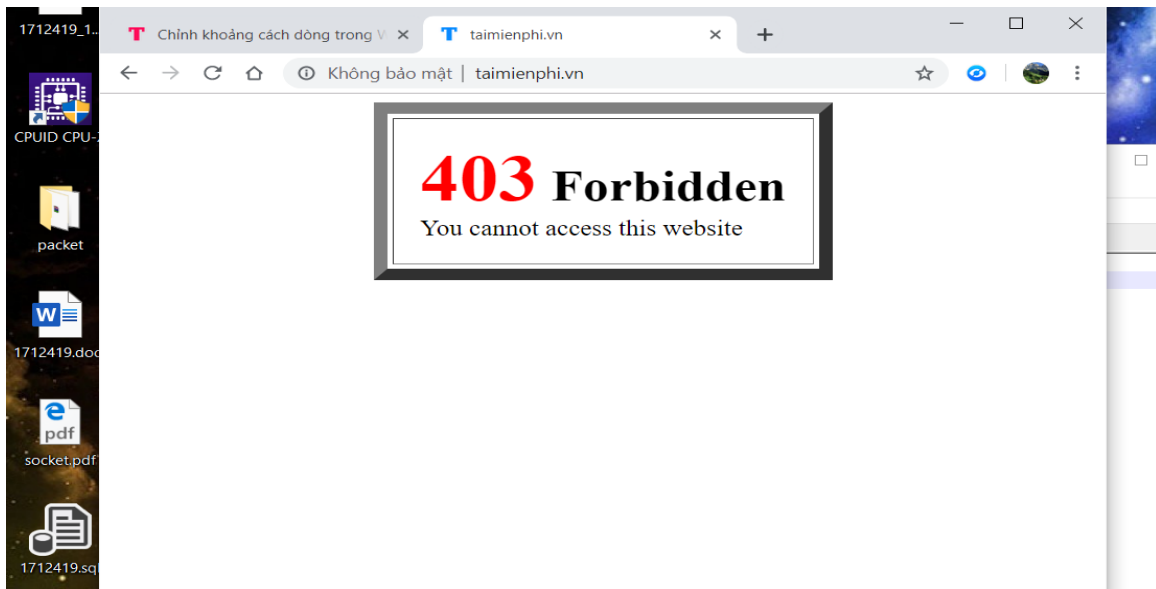
Tiến hành load trang web và trang web **chưa có** trong blacklist.



Tiến hành load trang web và trang web **có** trong blacklist.



403 Forbidden.



5. Đánh giá mức độ hoàn thành

Yêu cầu	Đánh giá
Hỗ trợ HTTP 1.0 và HTTP 1.1.	100%
Chương trình cho phép Client truy cập website thông qua Proxy Server (GET, POST).	100%
Chương trình chỉ cần hỗ trợ http, không cần https.	100%
Proxy Server phải xử lý đồng thời được các request từ client.	100%
Proxy Server sẽ chạy trên port 8888.	100%
Proxy Server cấu hình 1 file blacklist.conf, mỗi dòng chứa các domain website cấm Client truy cập. Proxy Server phải chặn tất cả các truy cập trùng với các domain trong file.	100%
Proxy Server có chức năng caching lại những nội dung của trang web, để phục vụ cho Client khác nhanh hơn. Sinh viên tự đưa ra cơ chế caching.	30%

Đánh giá mức độ hoàn thành: 90%

Proxy Server có chức năng caching lại những nội dung của trang web, để phục vụ cho Client khác nhanh hơn. Sinh viên tự đưa ra cơ chế caching. Ở yêu cầu này thì nhóm mới chỉ lên ý tưởng về caching nhưng chưa thể đưa code vào bài làm.

6. Bắt gói tin và quá trình gửi nhận dữ liệu của proxy-sever.

-Thực hiện kết nối tới trang: <http://muasamcong.mpi.gov.vn/>

- Thiết lập kết nối tới sever.

The screenshot displays the Wireshark interface with a packet capture of a web session. The left pane shows the packet list and packet details. The right pane shows the packet bytes. The selected packet is a TCP segment (No. 66) from 10.80.255.9 to 118.69.214.68, port 80, with sequence number 14722. The packet details show the TCP header and the application/javascript content type.

Packet List:

No.	Time	Source	Destination	Protocol	Length	Info
24	4.518001	10.80.255.9	118.69.214.68	TCP	66	14722 → 80 [SYN] Seq=0 Win=64240 Len=0 MSS=1460 WS=256 SACK_PERM=1
25	4.541801	118.69.214.68	10.80.255.9	TCP	66	80 → 14721 [SYN, ACK] Seq=0 Ack=1 Win=14600 Len=0 MSS=1412 SACK_PERM=1 WS=256
26	4.542236	10.80.255.9	118.69.214.68	TCP	54	14721 → 80 [ACK] Seq=1 Ack=1 Win=66304 Len=0
27	4.566843	118.69.214.68	10.80.255.9	TCP	66	80 → 14722 [SYN, ACK] Seq=0 Ack=1 Win=14600 Len=0 MSS=1412 SACK_PERM=1 WS=256
28	4.567080	10.80.255.9	118.69.214.68	TCP	54	14722 → 80 [ACK] Seq=1 Ack=1 Win=66304 Len=0
36	6.952736	10.80.255.9	118.69.214.68	HTTP	1253	POST /countersite/counter.php HTTP/1.1 (application/javascript)
37	6.976620	10.80.255.9	118.69.214.68	HTTP	1163	GET / HTTP/1.1
38	6.983397	118.69.214.68	10.80.255.9	TCP	60	80 → 14721 [ACK] Seq=1 Ack=1200 Win=17024 Len=0
39	6.987978	118.69.214.68	10.80.255.9	HTTP	368	HTTP/1.1 200 OK (text/html)
40	7.003202	118.69.214.68	10.80.255.9	TCP	60	80 → 14722 [ACK] Seq=1 Ack=1110 Win=16896 Len=0
41	7.028478	10.80.255.9	118.69.214.68	TCP	54	14721 → 80 [ACK] Seq=1200 Ack=315 Win=66048 Len=0
42	7.338508	118.69.214.68	10.80.255.9	TCP	1466	80 → 14722 [ACK] Seq=1 Ack=1110 Win=16896 Len=1412 [TCP segment of a reassembled data stream]
43	7.338511	118.69.214.68	10.80.255.9	TCP	1466	80 → 14722 [ACK] Seq=1413 Ack=1110 Win=16896 Len=1412 [TCP segment of a reassembled data stream]
44	7.338513	118.69.214.68	10.80.255.9	TCP	1466	80 → 14722 [ACK] Seq=2825 Ack=1110 Win=16896 Len=1412 [TCP segment of a reassembled data stream]
45	7.338522	118.69.214.68	10.80.255.9	TCP	1466	80 → 14722 [ACK] Seq=4237 Ack=1110 Win=16896 Len=1412 [TCP segment of a reassembled data stream]

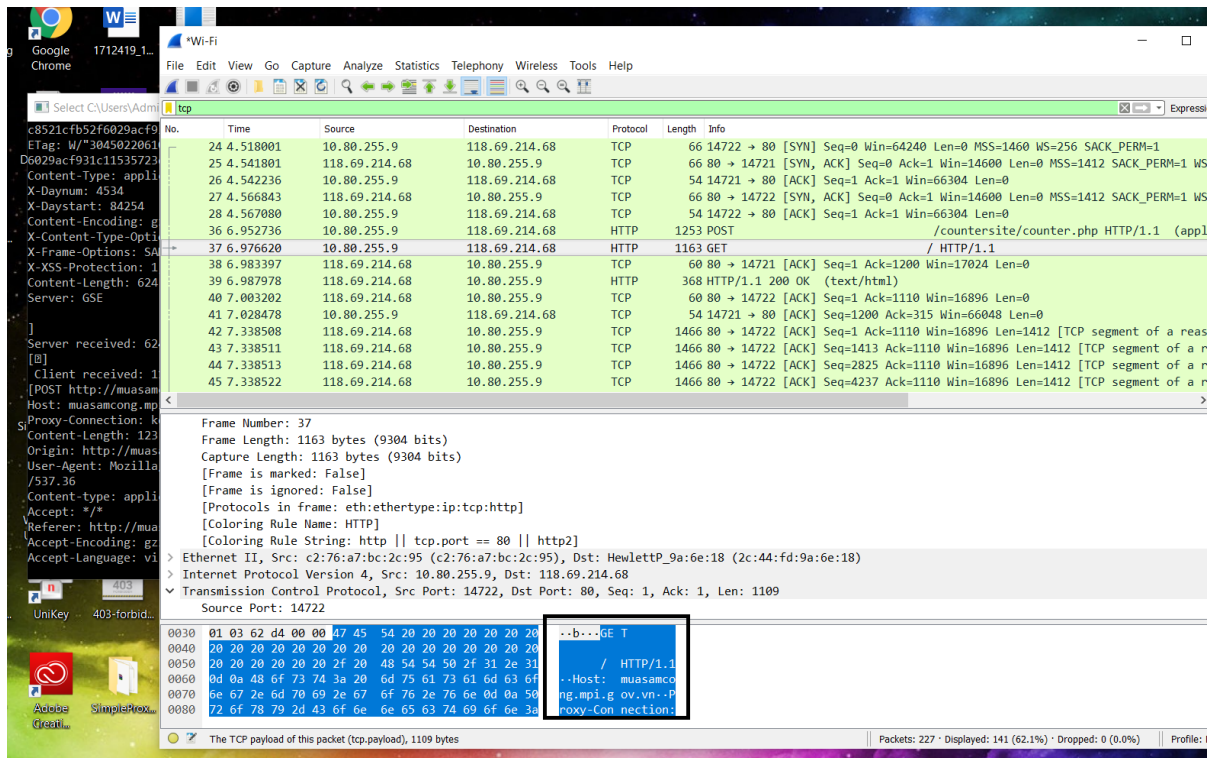
Packet Details:

Frame 24: 66 bytes on wire (528 bits), 66 bytes captured (528 bits) on interface 0
Ethernet II, Src: c2:76:a7:bc:2c:95 (c2:76:a7:bc:2c:95), Dst: HewlettP_9a:6e:18 (2c:44:fd:9a:6e:18)
Internet Protocol Version 4, Src: 10.80.255.9, Dst: 118.69.214.68
Transmission Control Protocol, Src Port: 14722, Dst Port: 80, Seq: 0, Len: 0
Source Port: 14722
Destination Port: 80
[Stream index: 2]
[TCP Segment Len: 0]
Sequence number: 0 (relative sequence number)
[Next sequence number: 0 (relative sequence number)]
Acknowledgment number: 0
1000 = Header Length: 32 bytes (8)

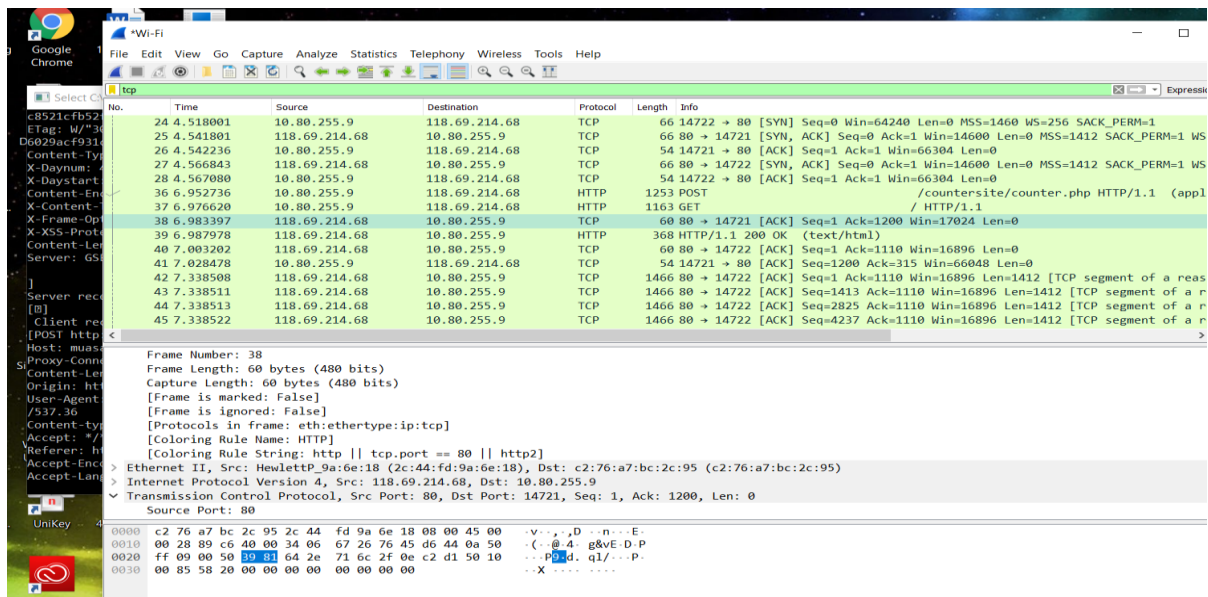
Packet Bytes:

```
0000 2c 44 fd 9a 6e 18 02 76 a7 bc 2c 95 00 00 45 00  ,D-n-v,,-E-
0010 00 34 3f e6 40 00 00 06 64 fa 0a 50 ff 09 76 45  4?@...d..P..vE
0020 d6 44 39 82 00 50 82 ff 94 04 00 00 00 00 80 02  D9..P..-e-
0030 fa f0 cd 65 00 00 02 04 05 b4 01 03 03 08 01 01  -e-
0040 04 02
```

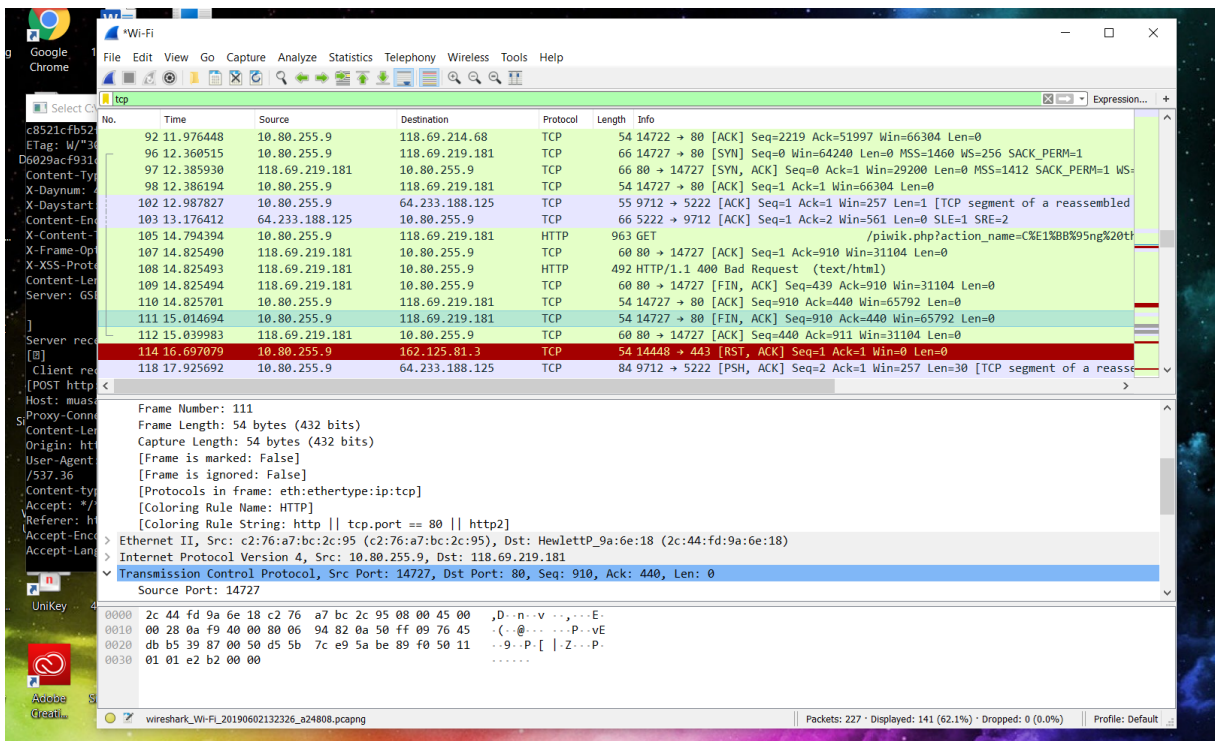
- Gửi yêu cầu kết nối:



- Server sẽ gửi tín hiệu gói tin đến người dùng và khi nhận thành công sẽ gửi tín hiệu phản hồi.



- Server sẽ gửi tín hiệu cờ FIN để thông báo kết thúc, khi người dùng bắt được cờ FIN sẽ gửi phản hồi lại xác nhận kết thúc.



7. Trong thực tế, tại sao lại cần Proxy Server?

- Đối với các nhà cung cấp dịch vụ đường truyền internet: Do internet có nhiều lượng thông tin mà theo quan điểm của từng quốc gia, từng chủng tộc hay địa phương mà các nhà cung cấp dịch vụ internet khu vực đó sẽ phối hợp sử dụng proxy với kỹ thuật tường lửa để tạo ra một bộ lọc gọi là firewall proxy nhằm ngăn chặn các thông tin độc hại hoặc trái thuần phong mỹ tục đối với quốc gia, chủng tộc hay địa phương đó. Địa chỉ các website mà khách hàng yêu cầu truy cập sẽ được lọc tại bộ lọc này, nếu địa chỉ không bị cấm thì yêu cầu của khách hàng tiếp tục được gửi đi, tới các DNS server của các nhà cung cấp dịch vụ. Firewall proxy sẽ lọc tất cả các thông tin từ internet gửi vào máy của khách hàng và ngược lại.

- 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.
- 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
- 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.
- 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.
- 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.

- 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.

Chúng ta đã thấy những ưu điểm của proxy server ,vậy proxy server có khuyết điểm gì không ?

- Khuyết điểm đầu tiên mà chúng ta thấy do không truy xuất trực tiếp ra bên ngoài mà phải không qua một proxy Server nên tốc độ truy xuất chậm hơn so với thực tế .
- Không phải lúc nào chúng ta cũng có thể kiểm một proxy server còn sống (alive) để sử dụng .Và nguy cơ có thể bị tấn công nếu proxy server đó mang thông tin nguy hiểm như hacker lập ra để phục vụ mục đích xấu.
- Đôi khi cần một proxy khác nhau cho mỗi nghi thức, bởi vì proxy server phải hiểu nghi thức đó để xác định những gì được phép và không được phép. Để thực hiện nhiệm vụ như là client đến server thật và server thật đến proxy client, sự kết hợp, install và config tất cả những server khác nhau đó có thể rất khó khăn .
- Mặc dù phần mềm proxy có hiệu quả rộng rãi những dịch vụ lâu đời và đơn giản như FPT và Telnet, những phần mềm mới và ít được sử dụng rộng rãi thì hiếm khi thấy. Thường đó chính là sự chậm trễ giữa thời gian xuất hiện một dịch vụ mới và proxy cho dịch vụ đó, khoảng thời gian phụ thuộc vào phương pháp thiết kế proxy cho dịch vụ đó, điều này cho thấy khá khó khăn khi đưa dịch vụ mới vào hệ thống khi chưa có proxy cho nó thì nên đặt bên ngoài fire wall, bởi vì nếu đặt bên trong hệ thống thì đó chính là yếu điểm.
- Nếu chúng ta “chịu khó” bỏ qua những khuyết điểm của Proxy Server ,và những hiệu quả của Proxy Server mang lại thì chúng ta sẽ có một lá chắn tốt cho hệ thống .

8. Ngôn ngữ lập trình sử dụng và Công cụ hỗ trợ.

Ngôn ngữ: C++

Công cụ hỗ trợ:

- IDE: Visual Studio 2013
- WireShark
- Google Chrome

9. Tài liệu tham khảo

1. <https://www.w3.org/Protocols/rfc2616/rfc2616.txt>
2. <https://github.com/>
3. <https://www.ntu.edu.sg/home/ehchua/programming/>
4. <https://www.howkteam.vn/>
5. <https://www.youtube.com/>