BÀI THỰC HÀNH SỐ 5

**PHÂN TÍCH HOẠT ĐỘNG CỦA GIAO THỨC DNS VÀ HTTP**

**1. Mục đích và yêu cầu**

**1.1. Mục đích**

Bài thí nghiệm này được thiết kế để trang bị cho sinh viên các kỹ năng sử dụng phần mềm Wireshark để bắt và lọc các gói tin DNS, HTTP theo yêu cầu. Thông qua đó, sinh viên có thể quan sát và hiểu được các hoạt động quan trọng của hai giao thức này. Bên cạnh đó, thông qua việc vận dụng kiến thức lý thuyết, sinh viên có thể giải thích kết quả đã quan sát được.

**1.2. Yêu cầu đối với sinh viên**

* Môi trường thực hành:
* Sử dụng thành thạo các chức năng cơ bản của phần mềm Wireshark
* Thực hiện thành thạo các thao tác trên hệ điều hành Windows, bao gồm các thao tác với thông số TCP/IP đã được hướng dẫn trong các bài thực hành trước.
* Kiến thức: Nắm vững kiến thức về tầng ứng dụng, các giao thức DNS và HTTP.
* Viết báo cáo thực hành và nộp kết quả theo yêu cầu như sau:
* Báo cáo(bản giấy) theo mẫu đã cung cấp
* File lưu lượng lab05.pcapng đặt trong thư mục có tên định dạng **TenSV\_ MSSV\_Lab05**. Nén thư mục và gửi vào địa chỉ email theo yêu cầu của cán bộ hướng dẫn thực hành.

**2. Cơ sở lý thuyết**

**2.1. Tên miền và hệ thống DNS**

Tên miền là một chuỗi ký tự định danh cho nút mạng, bao gồm các nhãn (label) cách nhau bởi dấu ‘.’, ví dụ soict.hust.edu.vn là tên miền cho máy chủ Web của Viện CNTT-TT, Đại học Bách khoa Hà Nội. Đối với người dùng, thay vì phải nhớ địa chỉ IP là các giá trị số khó nhớ, người dùng có thể truy cập vào dịch vụ bằng tên miền của máy chủ. Tuy nhiên, trong quá trình truyền tin, các nút mạng lại sử dụng địa chỉ IP. Do đó, tên miền cần phải được ánh xạ tới một hoặc nhiều địa chỉ IP nào đó. Trên mạng Internet, tên miền và ánh xạ này, cùng với các thông tin khác, được quản lý bởi các máy chủ trong hệ thống tên miền DNS. Bên cạnh đó, các máy chủ DNS này cung cấp dịch vụ tìm kiếm thông tin tên miền. Khi một nút mạng muốn gửi thông tin tới nút mạng khác mà chỉ biết tên miền, nó sẽ phải thực hiện quá trình tìm kiếm thông tin tên miền. Trong hầu hết các trường hợp, quá trình này bắt đầu bằng việc client gửi thông điệp DNS Query yêu cầu truy vấn tới máy chủ DNS. Kết quả tìm kiếm được đóng gói trong thông điệp DNS Response và trả lại cho client. Các bên trong dịch vụ DNS sử dụng giao thức UDP của tầng giao vận để truyền thông điệp với số hiệu cổng dịch vụ chuẩn trên máy chủ là 53.

**2.2. Dịch vụ Web và giao thức HTTP**

World Wide Web, gọi tắt là Web, lần đầu tiên được giới thiệu bởi Tim Berners-Lee vào năm 1991 với ý tưởng chính là liên kết các thông tin trên mạng Internet qua địa chỉ URL (Uniform Resource Location) và trình bày thành một văn bản sử dụng mã HTML (Hyper Text Markup Language) gọi là Webpage. Tập hợp các Webpage được lưu trữ trên một máy chủ Web để tạo thành một Website. Người dùng có thể dễ dàng sử dụng trình duyệt Web như là một phần mềm client để truy cập vào Website. Mặc dù là dịch vụ ra đời muộn hơn so với các dịch vụ truyền thống khác trên Internet, như là email hay truyền file, nhưng nhờ sự dễ dàng trong việc liên kết và chia sẻ thông tin mà Web đã nhanh chóng phổ biến và phát triển với tốc độ chóng mặt. Cho đến ngày nay, Web vừa là dịch vụ phổ biến nhất trên mạng Internet, vừa là một nền tảng để phát triển các dịch vụ khác.

Giao thức HTTP được Tim Berners-Lee phát triển để điều khiển hoạt động của dịch vụ Web. So với phiên bản HTTP 0.9 và HTTP 1.0 ở giai đoạn trước, phiên bản HTTP 1.1 hiện nay đã có nhiều cải tiến để nâng cao hiệu năng hoạt động của dịch vụ. Tuy nhiên, các nguyên lý cơ bản trong hoạt động vẫn được giữ nguyên. HTTP là một giao thức hướng liên kết, trong đó nó sử dụng dịch vụ của giao thức TCP trên tầng giao vận để thiết lập liên kết và điều khiển truyền các thông điệp HTTP trên liên kết đó. Máy chủ Web sử dụng cổng dịch vụ có số hiệu 80 để lắng nghe các yêu cầu thiết lập liên kết được gửi tới từ client. Để yêu cầu nội dung của Website, clien gửi đi thông điệp HTTP Request và chờ nhận thông điệp HTTP Response trả lời. Hiện nay, do các vấn đề về bảo mật, giao thức HTTPS dần được thay thế để đảm bảo an toàn cho quá trình truyền tin trong dịch vụ Web. HTTPS là cải tiến của HTTP, trong đó liên kết SSL/TLS được sử dụng thay cho lên kết TCP và số hiệu cổng ứng dụn là 443. Trên liên kết SSL/TLS, các thông điệp HTTP sẽ được mã hóa nhằm bảo vệ tính bí mật, toàn vẹn cho dữ liệu.

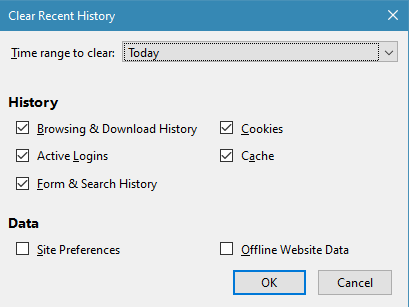
**3. Nội dung thực hành**

**3.1. Xác định thông số của máy trạm**

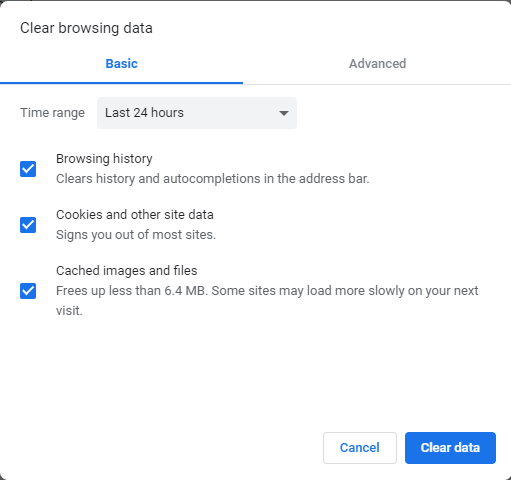
Sinh viên xác định địa chỉ IP, địa chỉ DNS Server trên máy tính ở phòng thực hành và ghi vào báo cáo. Để có được thông tin này, sinh viên xem lại bài thực hành số 2 và 3.

**3.2. Thu thập lưu lượng mạng**

* **Bước 1:** Tắt các chương trình của người dùng có trao đổi dữ liệu trên mạng trừ trình duyệt Web.
* **Bước 2:** Xóa bộ đệm của trình duyệt
* Mozilla Firefox: Nhấn tổ hợp phím Ctrl + Shift + Del. Chọn các mục như dưới đây và nhấn OK.



* Google Chrome: Nhấn tổ hợp phím Ctrl + Shift + Del. Chọn the past day. Chọn Cached images and files. Nhấp nút Clear data.



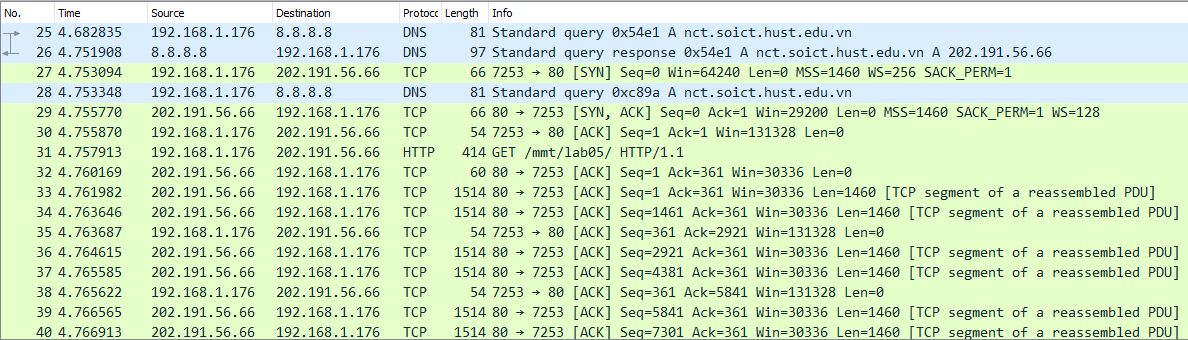
* **Bước 3:** Trên cửa sổ Command Prompt, thực hiện lệnh ipconfig /flushdns
* **Bước 4:** Khởi động phần mềm Wireshark và chọn bắt gói tin trên cạc mạng phù hợp
* **Bước 5:** Trên trình duyệt Web, mở cửa sổ duyệt riêng tư (Private Browsing):
* **Mozilla Firefox:** Nhấn tổ hợp phím Ctrl + Shift + P
* **Google Chrome:** Nhấn tổ hợp phím Ctrl + Shift + N

Truy cập vào địa chỉ sau:

http://nct.soict.hust.edu.vn/mmt/lab05/

*Lưu ý: Các bước trên có tác dụng giúp ta quan sát đầy đủ hơn hoạt động của các dịch vụ.*

* **Bước 6:** Sau khi trình duyệt đã tải xong trang Web khoảng 5-10 giây, dừng việc bắt gói tin trên Wireshark. Hình ảnh lưu lượng bắt được trên Wireshark có thể như sau:



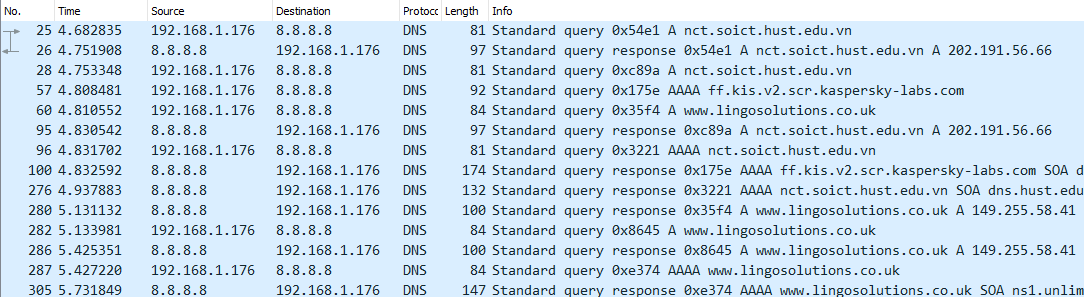
***Lưu ý:***

* *Trên menu của Wireshark chọn Analyze 🡪 Enabled Protocols. Kiểm tra để chắc chắn các mục DNS, HTTP đã được chọn.*
* *Nếu file lưu lượng trên máy sinh viên không có các gói tin có Protocol là DNS thì thực hiện lại từ bước 2.*
* *Các gói tin bắt được trên máy sinh viên có thể sẽ có một số thông số khác với hình ảnh minh họa. Điều này là hoàn toàn bình thường và không có ảnh hưởng tới quá trình thực hành*
* **Bước 7:** Lưu file lưu lượng có tên là **lab05.pcapng** và nộp cùng báo cáo thực hành

**3.3. Quan sát quá trình truyền dữ liệu trong DNS**

Sử dụng file lưu lượng ở mục 3.2 để quan sát và trả lời các câu hỏi.

* **Bước 1:** Điền giá trị **dns** vào mục Filter của Wireshark để lọc ra các thông diệp DNS đã bắt được tương tự như hình minh họa dưới đây.



* **Bước 2:** Chọn gói tin DNS đượctrình duyệt gửi đi để yêu cầu phân giải tên miền của Website đã truy cập và trả lời câu hỏi 1.

*Lưu ý: Nếu không tìm thấy gói tin nào, sinh viên cần thực hiện lại các thao tác của mục 3.2.*

|  |
| --- |
| **Câu hỏi 1(*1 điểm*):** Hãy xác định các thông tin sau trên thông điệp   * *STT gói tin(No.):* * *Giao thức tầng giao vận được sử dụng để gửi thông điệp đi:* * *Địa chỉ IP nguồn:* * *Số hiệu cổng ứng dụng nguồn:* * *Địa chỉ IP đích:* * *Số hiệu cổng đích? Đây là số hiệu cổng ứng dụng của dịch vụ nào?* * *Kiểu thông tin truy vấn(Type):*   Qua việc xác định các thông số mạng trên máy trạm của sinh viên ở mục 3.1, cho biết thông điệp này được gửi tới nút mạng nào? |

* **Bước 3**: Tìm thông điệp DNS Response trả lời cho thông điệp yêu cầu ở bước 2 để quan sát và trả lời câu hỏi 2

|  |
| --- |
| **Câu hỏi 2(*1 điểm*):** Hãy xác định các thông tin sau trên thông điệp   * *STT gói tin(No.):* * *Giao thức tầng giao vận được sử dụng để gửi thông điệp đi:* * *Địa chỉ IP nguồn:* * *Số hiệu cổng ứng dụng nguồn:* * *Địa chỉ IP đích:* * *Số hiệu cổng đích:* * *Kiểu thông tin truy vấn(Type):* * *Tên miền được truy vấn:* * *Địa chỉ IP của tên miền được truy vấn:*   Tại sao xác định được đây là thông điệp trả lời cho thông điệp yêu cầu ở bước 2? |

* **Bước 4:** Quan sát tất cả các thông điệp DNS và trả lời câu hỏi 3?

|  |
| --- |
| **Câu hỏi 3(*1 điểm*):** Tại sao ngoài tên miền nct.soict.hust.edu.vn được truy vấn do người dùng truy cập vào trang Web http://nct.soict.hust.edu.vn/mmt/lab05/, còn có truy vấn tới tên miền khác. Các tên miền khác được truy vấn và địa chỉ IP của các tên miền đó là gì? |

**3.4. Quan sát quá trình truyền dữ liệu của HTTP**

Sử dụng file lưu lượng ở mục 3.2 để quan sát và trả lời các câu hỏi.

* **Bước 1:** Giả sử ở phần trên, sinh viên đã quan sát được địa chỉ IP phân giải từ tên miền nct.soict.hust.edu.vn của máy chủ Web là X. Điền giá trị **ip.addr == X** *(Lưu ý: Thay X bằng địa chỉ IP đã quan sát được)* vào mục Filter của Wireshark. Sinh viên sẽ quan sát thấy các thông điệp mà máy trạm trao đổi với máy chủ Web.

|  |
| --- |
| **Câu hỏi 4(*1 điểm*):** Trước khi thông điệp HTTP đầu tiên được gửi đi tới máy chủ nct.soict.hust.edu.vn, máy trạm và máy chủ đã thực hiện quá trình gì? Số thứ tự (No) của các gói tin trong quá trình đó mà sinh viên quan sát được là gì? Số hiệu cổng ứng dụng của các bên đã sử dụng là bao nhiêu? Số hiệu cổng ứng dụng trên máy chủ là cổng ứng dụng của dịch vụ nào? |

* **Bước 2:** Điền giá trị **http** vào mục Filter của Wireshark. Sinh viên sẽ quan sát thấy các gói tin HTTP mà máy trạm đã trao đổi trên mạng.
* **Bước 3:** Quan sát các thông điệp HTTP trao đổi giữa máy trạm và máy chủ Web nct.soict.hust.edu.vn và trả lời câu hỏi 5

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Câu hỏi 5(2 *điểm*):** Có bao nhiêu thông điệp HTTP Request được gửi đi? Liệt kê các thông tin sau về các thông điệp HTTP giữa máy trạm và máy chủ Web nct.soict.hust.edu.vn   |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | HTTP Request | | | HTTP Response | | | | No. | Phương thức yêu cầu | Đối tượng yêu cầu | No. | Mã trả lời | Ý nghĩa mã trả lời | |  |  |  |  |  |  |   Trong các thông điệp HTTP Request, có những thông điệp nào được gửi đi liên tiếp mà không đợi thông điệp trả lời từ máy chủ không? Nếu có, tại sao trình duyệt Web trên máy trạm thực hiện như vậy? |

* **Bước 4**: Chọn thông điệp HTTP Request đầu tiên được máy trạm gửi cho máy chủ Web nct.hust.edu.vn và trả lời câu hỏi 6.

|  |
| --- |
| **Câu hỏi 6(*1 điểm*):** Hãy cho biết các thông tin sau về thông điệp yêu cầu:   * *Giao thức tầng giao vận được sử dụng để truyền thông điệp* * *Số hiệu cổng ứng dụng đích* * *Phiên bản của giao thức HTTP mà máy trạm sử dụng* * *Giá trị của trường Connection trong tiêu đề HTTP* |

* **Bước 5:** Tìm thông điệp HTTP Response mà máy chủ Web trả lời cho thông điệp yêu cầu ở bước 5 và trả lời câu hỏi 7

|  |
| --- |
| **Câu hỏi 7(*1 điểm*):** Hãy cho biết các thông tin sau về thông điệp trả lời:   * *Phiên bản của giao thức HTTP mà máy chủ sử dụng* * *Giá trị của trường Connection trong tiêu đề HTTP* * *Phần thân chứa dữ liệu gì? Dữ liệu này có kích thước là bao nhiêu?* * *Thông điệp này đóng gói trong bao nhiêu gói tin TCP?*   Sau khi thông điệp này được gửi đi, kết nối TCP còn được duy trì không? |

|  |
| --- |
| **Câu hỏi 8(*1 điểm*):** Ngoài quá trình trao đổi dữ liệu với máy chủ Web nct.soict.hust.edu.vn, máy trạm còn gửi thông điệp HTTP Request tới máy chủ Web có tên miền và địa chỉ IP là gì? Tại sao máy trạm phát đi thông điệp này? *(Gợi ý: Xem lại nội dung trong phần thân của thông điệp HTTP Response trong bước 5)*  Xem phần tiêu đề của thông điệp HTTP Request trên và cho biết giá trị trường **Referer** là gì? |

|  |
| --- |
| **Câu hỏi 9(*1 điểm*):** Đoạn sau đây mô tả ngắn gọn quá trình xử lý truy cập vào một trang Web trên trình duyệt Web. Hãy điền vào chỗ trống cụm từ còn thiếu  Khi nhận được yêu cầu truy cập vào một trang Web nào đó qua địa chỉ URL, nếu chưa biết địa chỉ IP của máy chủ Web. Trình duyệt gửi thông điệp DNS (Domain Name System) tới máy chủ DNS Trong thông điệp DNS trả lời nhận được, trình duyệt xác định được địa chỉ IP của máy chủ Web. Sau đó, trình duyệt gửi yêu cầu để thiết lập kết nối TCP/IP với máy chủ Web. Trên kết nối đã được thiết lập, trình duyệt gửi đi thông điệp HTTP (HyperText Transfer Protocol) để yêu cầu nội dung của trang Web. Máy chủ Web tìm kiếm nội dung được yêu cầu và trả lại thông điệp HTTP cùng với mã trả lời 200 nếu tìm thấy, hoặc mã mã lỗi (ví dụ: 404) nếu không tìm thấy. Nếu hai bên sử dụng giao thức HTTP có phiên bản liên kết (ví dụ: HTTP/1.1) thì liên kết sẽ được duy trì cho tới khi trình duyệt đã tải xong nội dung trang Web từ máy chủ. |