**TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN**
**KHOA CÔNG NGHỆ THÔNG TIN**
**BỘ MÔN MẠNG MÁY TÍNH & VIỄN THÔNG**

**PHAN TRUNG HIẾU - TRẦN LÊ QUÂN**

# CÁC PHƯƠNG PHÁP LẬP TRÌNH VƯỢT FIREWALL

**KHÓA LUẬN CỬ NHÂN TIN HỌC**



**NIÊN KHÓA 2001 - 2005**

**TRƯỜNG ĐẠI HỌC KHOA HỌC TỰ NHIÊN**
**KHOA CÔNG NGHỆ THÔNG TIN**
**BỘ MÔN MẠNG MÁY TÍNH & VIỄN THÔNG**

**PHAN TRUNG HIẾU 0112463**
**TRẦN LÊ QUÂN 0112319**

# CÁC PHƯƠNG PHÁP LẬP TRÌNH VƯỢT FIREWALL



**KHÓA LUẬN CỬ NHÂN TIN HỌC**

**GIÁO VIÊN HƯỚNG DẪN**
**Th.S ĐỖ HOÀNG CƯỜNG**

**NIÊN KHÓA 2001 – 2005**

# LỜI NHẬN XÉT CỦA GIÁO VIÊN HƯỚNG DẪN

# LỜI NHẬN XÉT CỦA GIÁO VIÊN PHẢN BIỆN

# LỜI CẢM ƠN

Sau hơn 6 tháng nỗ lực thực hiện, luận văn nghiên cứu "Các phương pháp lập trình vượt firewall" đã phần nào hoàn thành. Ngoài sự nỗ lực của bản thân, chúng em đã nhận được sự khích lệ rất nhiều từ phía nhà trường, thầy cô, gia đình và bạn bè trong khoa. Chính điều này đã mang lại cho chúng em sự động viên rất lớn để chúng em có thể hoàn thành tốt luận văn của mình.

Trước hết, chúng con xin cảm ơn những bậc làm cha, làm mẹ đã luôn ủng hộ, chăm sóc chúng con và tạo mọi điều kiện tốt nhất để chúng con có thể hoàn thành nhiệm vụ của mình.

Chúng em xin cảm ơn nhà trường nói chung và Khoa CNTT nói riêng đã đem lại cho chúng em nguồn kiến thức vô cùng quý giá để chúng em có đủ kiến thức hoàn thành luận văn cũng như làm hành trang bước vào đời.

Em xin cảm ơn các thầy cô thuộc bộ môn MMT, đặc biệt là thầy Đỗ Hoàng Cường – giáo viên hướng dẫn của chúng em đã tận tình hướng dẫn và giúp đỡ chúng em mỗi khi chúng em có khó khăn trong quá trình học tập cũng như trong quá trình làm luận văn tốt nghiệp.

Xin cảm ơn tất cả các bạn bè thân yêu đã động viên, giúp đỡ chúng em trong suốt quá trình học tập cũng như làm đề tài.

Một lần nữa, xin cảm ơn tất cả mọi người…

TPHCM 7/2005

Nhóm sinh viên thực hiện

Phan Trung Hiếu – Trần Lê Quân

# LỜI NÓI ĐẦU

Nội dung luận văn được trình bày trong 8 chương thuộc về 5 phần khác nhau :

*Phần thứ nhất:* **CƠ SỞ LÝ THUYẾT**

- Chương 1: Giới thiệu về firewall
- Chương 2: Khái niệm proxy
- Chương 3: Các phương pháp lập trình vượt firewall

*Phần thứ hai:* **CÁC PHƯƠNG PHÁP LẬP TRÌNH VƯỢT FIREWALL**

- Chương 4: Vượt firewall bằng HTTP proxy Servers
- Chương 5: Vượt firewall bằng Web-based proxy

*Phần thứ ba:* **MODULE CHỐNG VƯỢT FIREWALL**

- Chương 6: Plug-in chống vượt firewall cho trình duyệt Internet Explorer
- Chương 7: Service chống vượt Firewall

*Phần thứ tư:* **TỔNG KẾT**

- Chương 8: Kết luận.

*Phần thứ năm:* **PHỤ LỤC**

# MỤC LỤC

# DANH SÁCH HÌNH

# DANH SÁCH BẢNG

**PHẦN THỨ NHẤT**

# CƠ SỞ LÝ THUYẾT

## Chương 1: GIỚI THIỆU VỀ FIREWALL

### 1.1 Đặt vấn đề:

Song song với việc xây dựng nền tảng về công nghệ thông tin, cũng như phát triển các ứng dụng máy tính trong sản xuất, kinh doanh, khoa học, giáo dục, xã hội,... thì việc bảo về những thành quả đó là một điều không thể thiếu. Sử dụng các bức tường lửa (Firewall) để bảo vệ mạng nội bộ (Intranet), tránh sự tấn công từ bên ngoài là một giải pháp hữu hiệu, đảm bảo được các yếu tố:

- An toàn cho sự hoạt động của toàn bộ hệ thống mạng
- Bảo mật cao trên nhiều phương diện
- Khả năng kiểm soát cao
- Đảm bảo tốc độ nhanh
- Mềm dẻo và dễ sử dụng
- Trong suốt với người sử dụng
- Đảm bảo kiến trúc mở

### 1.2 Nhu cầu bảo vệ thông tin:

#### 1.2.1 ***Nguyên nhân:***

Ngày nay, Internet, một kho tàng thông tin khổng lồ, phục vụ hữu hiệu trong sản xuất kinh doanh, đã trở thành đối tượng cho nhiều người tấn công với các mục đích khác nhau. Đôi khi, cũng chỉ đơn giản là để thử tài hoặc đùa bỡn với người khác.

Cùng với sự phát triển không ngừng của Internet và các dịch vụ trên Internet, số lượng các vụ tấn công trên Internet cũng tăng theo cấp số nhân. Trong khi các phương tiện thông tin đại chúng ngày càng nhắc nhiều đến Internet với những khả năng truy nhập thông tin dường như đến vô tận của nó, thì các tài liệu chuyên môn bắt đầu đề cập nhiều đến vấn đề bảo đảm và an toàn dữ liệu cho các máy tính được kết nối vào mạng Internet.

Theo số liệu của CERT (Computer Emegency Response Team), số lượng các vụ tấn công trên Internet được thông báo cho tổ chức này là ít hơn 200 vào năm 1989, khoảng 400 vào năm 1991, 1400 vào năm 1993, và 2241 vào năm 1994. Những vụ tấn công này nhằm vào tất cả các máy tính có mặt trên Internet, các máy tính của tất cả các công ty lớn như AT&T, IBM, các trường đại học, các cơ quan nhà nước, các tổ chức quân sự, nhà băng... Một số vụ tấn công có quy mô khổng lồ (có tới 100.000 máy tính bị tấn công). Hơn nữa, những con số này chỉ là phần nổi của tảng băng. Một phần rất lớn các vụ tấn công không được thông báo, vì nhiều lý do, trong đó có thể kể đến nỗi lo bị mất uy tín, hoặc đơn giản những người quản trị hệ thống không hề hay biết những cuộc tấn công nhằm vào hệ thống của họ.

Không chỉ số lượng các cuộc tấn công tăng lên nhanh chóng, mà các phương pháp tấn công cũng liên tục được hoàn thiện. Điều đó một phần do các nhân viên quản trị hệ thống được kết nối với Internet ngày càng đề cao cảnh giác. Cũng theo CERT, những cuộc tấn công thời kỳ 1988-1989 chủ yếu đoán tên người sử dụng-mật khẩu (UserID-password) hoặc sử dụng một số lỗi của các chương trình và hệ điều hành (security hole) làm vô hiệu hệ thống bảo vệ, tuy nhiên các cuộc tấn công vào thời gian gần đây bao gồm cả các thao tác như giả mạo địa chỉ IP, theo dõi thông tin truyền qua mạng, chiếm các phiên làm việc từ xa (telnet hoặc rlogin).

Nhu cầu bảo vệ thông tin trên Internet có thể chia thành ba loại gồm: Bảo vệ dữ liệu; Bảo vệ các tài nguyên sử dụng trên mạng và Bảo vệ danh tiếng của cơ quan.

### 1.2.2 ***Bảo vệ dữ liệu:***

Những thông tin lưu trữ trên hệ thống máy tính cần được bảo vệ do các yêu cầu sau:

- Bảo mật: Những thông tin có giá trị về kinh tế, quân sự, chính sách vv... cần được giữ kín.
- Tính toàn vẹn: Thông tin không bị mất mát hoặc sửa đổi, đánh tráo.
- Tính kịp thời: Yêu cầu truy nhập thông tin vào đúng thời điểm cần thiết.

Trong các yêu cầu này, thông thường yêu cầu về bảo mật được coi là yêu cầu số 1 đối với thông tin lưu trữ trên mạng. Tuy nhiên, ngay cả khi những thông tin này không được giữ bí mật, thì những yêu cầu về tính toàn vẹn cũng rất quan trọng. Không một cá nhân, một tổ chức nào lãng phí tài nguyên vật chất và thời gian để lưu trữ những thông tin mà không biết về tính đúng đắn của những thông tin đó.

### 1.2.3 ***Bảo vệ các tài nguyên sử dụng trên mạng:***

Trên thực tế, trong các cuộc tấn công trên Internet, kẻ tấn công, sau khi đã làm chủ được hệ thống bên trong, có thể sử dụng các máy này để phục vụ cho mục đích của mình nhằm chạy các chương trình dò mật khẩu người sử dụng, sử dụng các liên kết mạng sẵn có để tiếp tục tấn công các hệ thống khác vv...

### 1.2.4 ***Bảo vệ danh tiếng cơ quan:***

Một phần lớn các cuộc tấn công không được thông báo rộng rãi, và một trong những nguyên nhân là nỗi lo bị mất uy tín của cơ quan, đặc biệt là các công ty lớn và các cơ quan quan trọng trong bộ máy nhà nước. Trong trường hợp người quản trị hệ thống chỉ được biết đến sau khi chính hệ thống của mình được dùng làm bàn đạp để tấn công các hệ thống khác, thì tổn thất về uy tín là rất lớn và có thể để lại hậu quả lâu dài.

## 1.3 Các kiểu tấn công:

### 1.3.1 ***Tấn công trực tiếp:***

Những cuộc tấn công trực tiếp thông thường được sử dụng trong giai đoạn đầu để chiếm được quyền truy nhập bên trong. Một phương pháp tấn công cổ điển là dò tìm tên ngời sử dụng và mật khẩu. Đây là phương pháp đơn giản, dễ thực hiện và không đòi hỏi một điều kiện đặc biệt nào để bắt đầu.

Kẻ tấn công có thể sử dụng những thông tin như tên người dùng, ngày sinh, địa chỉ, số nhà vv.. để đoán mật khẩu. Trong trường hợp có được danh sách người sử dụng và những thông tin về môi trường làm việc, có một trương trình tự động hoá về việc dò tìm mật khẩu này.

Một chương trình có thể dễ dàng lấy được từ Internet để giải các mật khẩu đã mã hoá của các hệ thống unix có tên là crack, có khả năng thử các tổ hợp các từ trong một từ điển lớn, theo những quy tắc do người dùng tự định nghĩa. Trong một số trường hợp, khả năng thành công của phương pháp này có thể lên tới 30%.

Phương pháp sử dụng các lỗi của chương trình ứng dụng và bản thân hệ điều hành đã được sử dụng từ những vụ tấn công đầu tiên và vẫn được tiếp tục để chiếm quyền truy nhập. Trong một số trường hợp phương pháp này cho phép kẻ tấn công có được quyền của người quản trị hệ thống (root hay administrator).

Hai ví dụ thường xuyên được đưa ra để minh hoạ cho phương pháp này là ví dụ với chương trình sendmail và chương trình rlogin của hệ điều hành UNIX.

Sendmail là một chương trình phức tạp, với mã nguồn bao gồm hàng ngàn dòng lệnh của ngôn ngữ C. **Sendmail được chạy với quyền ưu tiên của người quản trị hệ thống, do chương trình phải có quyền ghi vào hộp thư của những người sử dụng máy. Và Sendmail trực tiếp nhận các yêu cầu về thư tín trên mạng bên ngoài.** Đây chính là những yếu tố làm cho sendmail trở thành một nguồn cung cấp những lỗ hổng về bảo mật để truy nhập hệ thống.

Rlogin cho phép người sử dụng từ một máy trên mạng truy nhập từ xa vào một máy khác sử dụng tài nguyên của máy này. **Trong quá trình nhận tên và mật khẩu của người sử dụng, rlogin không kiểm tra độ dài của dòng nhập**, do đó kẻ tấn công có thể đưa vào một xâu đã được tính toán trước để ghi đè lên mã chương trình của rlogin, qua đó chiếm được quyền truy nhập.

### 1.3.2 *Nghe trộm:*

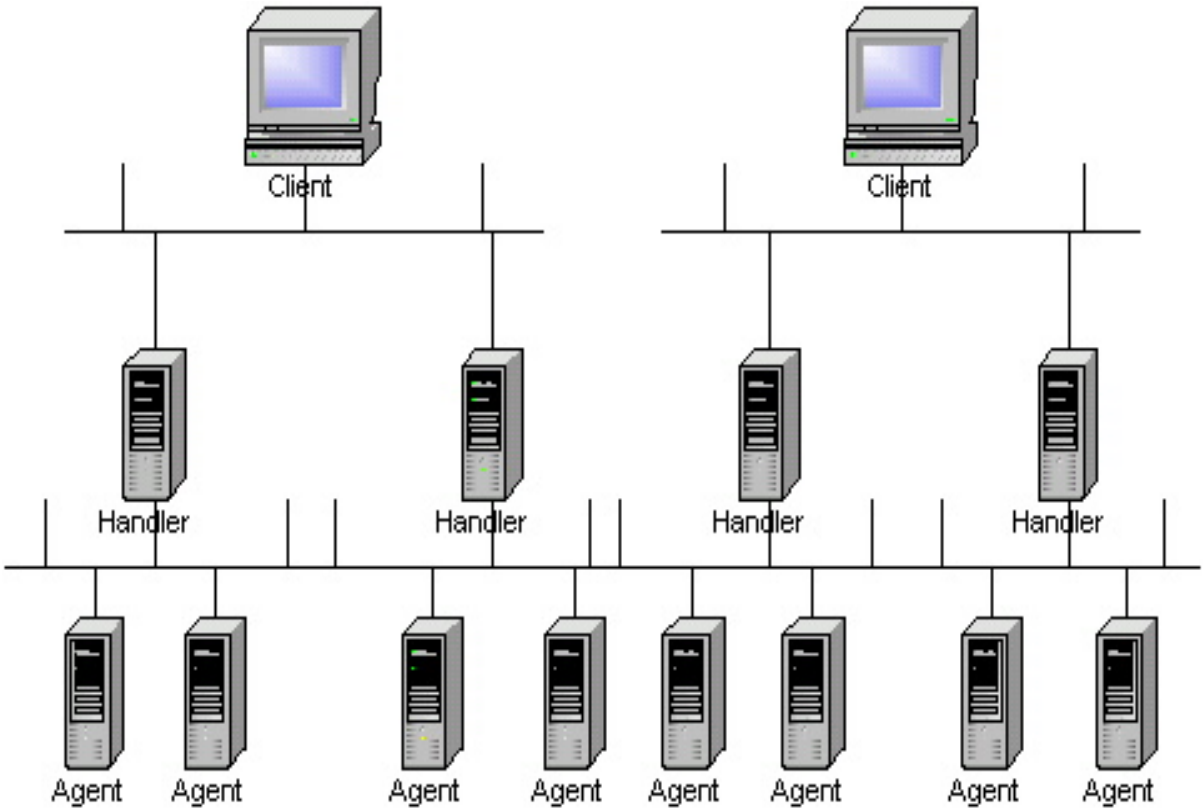
Việc nghe trộm thông tin trên mạng có thể đưa lại những thông tin có ích như tên, mật khẩu của người sử dụng, các thông tin mật chuyển qua mạng. **Việc nghe trộm thường được tiến hành ngay sau khi kẻ tấn công đã chiếm được quyền truy nhập hệ thống, thông qua các chương trình cho phép bắt các gói tin** vào chế độ nhận toàn bộ các thông tin lưu truyền trên mạng. Những thông tin này cũng có thể dễ dàng lấy được trên Internet.

### 1.3.3 *Giả mạo địa chỉ:*

Việc giả mạo địa chỉ IP có thể được thực hiện thông qua việc sử dụng khả năng dẫn đường trực tiếp (source-routing). Với cách tấn công này, kẻ tấn công gửi các gói tin IP tới mạng bên trong với một địa chỉ IP giả mạo (thông thường là địa chỉ của một mạng hoặc một máy được coi là an toàn đối với mạng bên trong), đồng thời chỉ rõ đường dẫn mà các gói tin IP phải gửi đi.

### 1.3.4 *Vô hiệu các chức năng của hệ thống (DoS, DDoS):*

Đây là kểu tấn công nhằm tê liệt hệ thống, không cho nó thực hiện chức năng mà nó thiết kế. **Kiểu tấn công này không thể ngăn chặn được**, do những phương tiện đợc tổ chức tấn công cũng chính là các phương tiện để làm việc và truy nhập thông tin trên mạng. Ví dụ sử dụng lệnh ping với tốc độ cao nhất có thể, buộc một hệ thống tiêu hao toàn bộ tốc độ tính toán và khả năng của mạng để trả lời các lệnh này, không còn các tài nguyên để thực hiện những công việc có ích khác.




**Hình 1 Mô hình tấn công DDoS**

- Client là một attacker sắp xếp một cuộc tấn công
- Handler là một host đã được thỏa hiệp để chạy những chương trình đặc biệt dùng đê tấn công
- Mỗi handler có khả năng điều khiển nhiều agent
- Mỗi agent có trách nhiệm gửi stream data tới victim

### 1.3.5 *Lỗi của người quản trị hệ thống:*

Đây không phải là một kiểu tấn công của những kẻ đột nhập, tuy nhiên lỗi của người quản trị hệ thống thường tạo ra những lỗ hổng cho phép kẻ tấn công sử dụng để truy nhập vào mạng nội bộ.

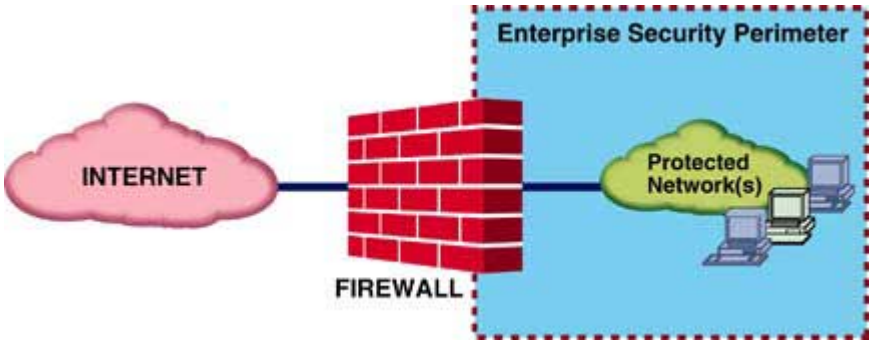
### 1.3.6 ***Tấn công vào yếu tố con người:***

Kẻ tấn công có thể liên lạc với một ngời quản trị hệ thống, giả làm một người sử dụng để yêu cầu thay đổi mật khẩu, thay đổi quyền truy nhập của mình đối với hệ thống, hoặc thậm chí thay đổi một số cấu hình của hệ thống để thực hiện các phương pháp tấn công khác. Với kiểu tấn công này không một thiết bị nào có thể ngăn chặn một cách hữu hiệu, và chỉ có một cách giáo dục người sử dụng mạng nội bộ về những yêu cầu bảo mật để đề cao cảnh giác với những hiện tượng đáng nghi. Nói chung yếu tố con ngời là một điểm yếu trong bất kỳ một hệ thống bảo vệ nào, và chỉ có sự giáo dục cộng với tinh thần hợp tác từ phía người sử dụng có thể nâng cao được độ an toàn của hệ thống bảo vệ.

## 1.4 Firewall là gì ?

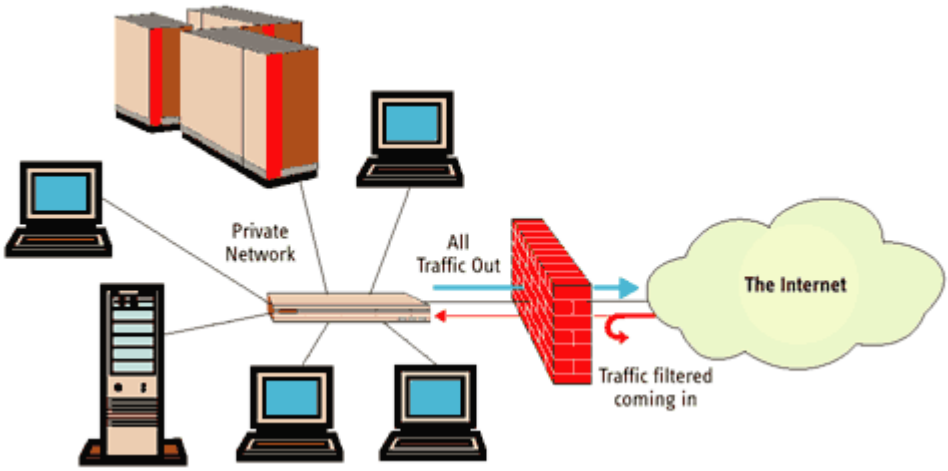
Thuật ngữ Firewall có nguồn gốc từ một kỹ thuật thiết kế trong xây dựng để ngăn chặn, hạn chế hoả hoạn. Trong công nghệ mạng thông tin, Firewall là một kỹ thuật được tích hợp vào hệ thống mạng để chống sự truy cập trái phép, nhằm bảo vệ các nguồn thông tin nội bộ và hạn chế sự xâm nhập không mong muốn vào hệ thống. Cũng có thể hiểu Firewall là một cơ chế (mechanism) để bảo vệ mạng tin tưởng (Trusted network) khỏi các mạng không tin tưởng (Untrusted network).

Thông thường Firewall được đặt giữa mạng bên trong (Intranet) của một công ty, tổ chức, ngành hay một quốc gia, và Internet. Vai trò chính là bảo mật thông tin, ngăn chặn sự truy nhập không mong muốn từ bên ngoài (Internet) và cấm truy nhập từ bên trong (Intranet) tới một số địa chỉ nhất định trên Internet.




**Hình 2 Mô hình firewall**

Một cách vắn tắt, firewall là hệ thống ngăn chặn việc truy nhập trái phép từ bên ngoài vào mạng cũng như những kết nối không hợp lệ từ bên trong ra. Firewall thực hiện việc lọc bỏ những địa chỉ không hợp lệ dựa theo các quy tắc hay chỉ tiêu định trước.




**Hình 3 Lọc gói tin tại firewall**

Firewall có thể là hệ thống phần cứng, phần mềm hoặc kết hợp cả hai. Nếu là phần cứng, nó có thể chỉ bao gồm duy nhất bộ lọc gói tin hoặc là thiết bị định tuyến (router được tích hợp sẵn chức năng lọc gói tin). Bộ định tuyến có các tính năng bảo mật cao cấp, trong đó có khả năng kiểm soát địa chỉ IP. Quy trình kiểm soát cho phép bạn định ra những địa chỉ IP có thể kết nối với mạng của bạn và ngược lại. Tính chất

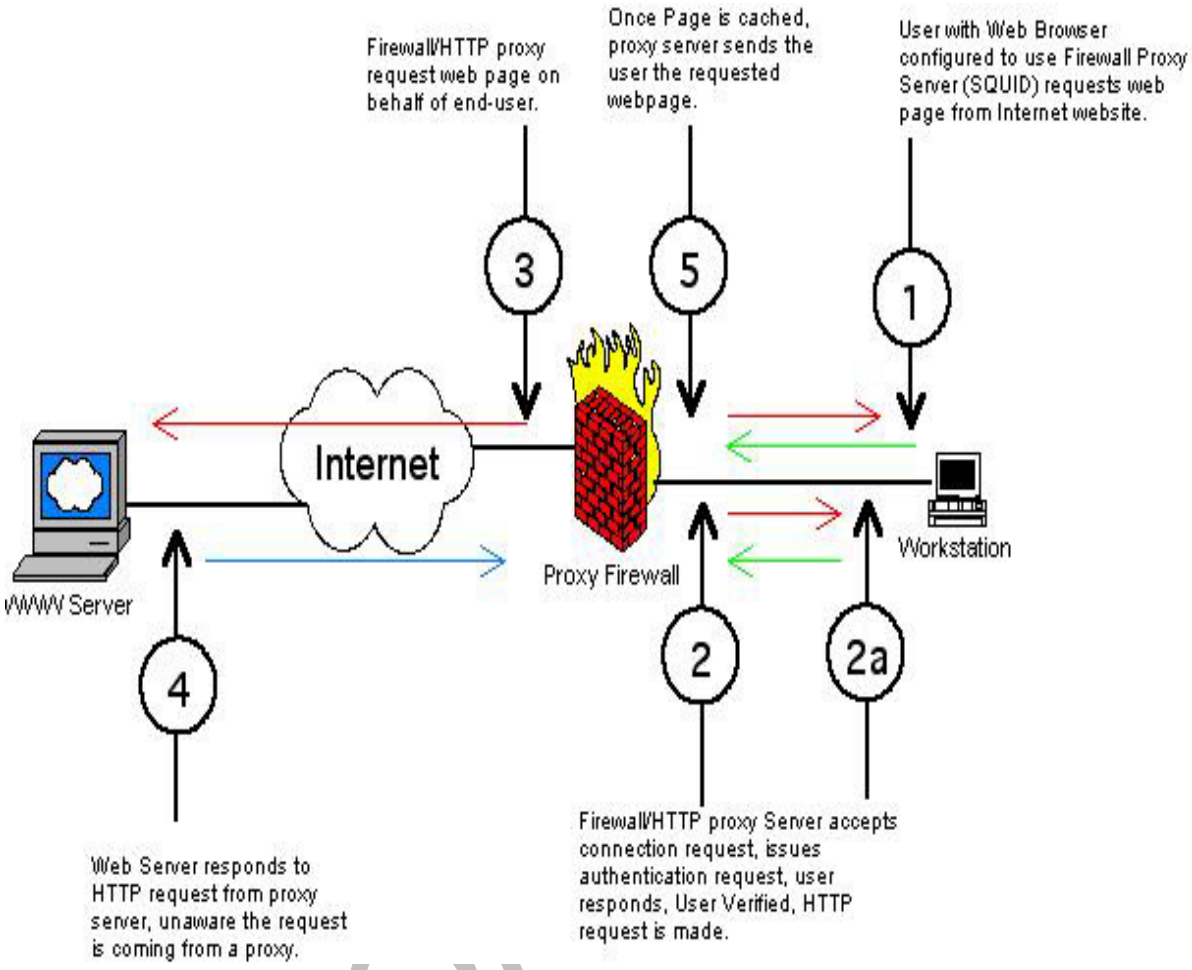
chung của các Firewall là phân biệt địa chỉ IP dựa trên các gói tin hay từ chối việc truy nhập bất hợp pháp căn cứ trên địa chỉ nguồn.

## 1.5 Các chức năng chính:

### 1.5.1 ***Chức năng:***

Chức năng chính của Firewall là kiểm soát luồng thông tin từ giữa Intranet và Internet. Thiết lập cơ chế điều khiển dòng thông tin giữa mạng bên trong (Intranet) và mạng Internet. Cụ thể là:

- Cho phép hoặc cấm những dịch vụ truy nhập ra ngoài (từ Intranet ra Internet).
- Cho phép hoặc cấm những dịch vụ phép truy nhập vào trong (từ Internet vào Intranet).
- Theo dõi luồng dữ liệu mạng giữa Internet và Intranet.
- Kiểm soát địa chỉ truy nhập, cấm địa chỉ truy nhập.
- Kiểm soát người sử dụng và việc truy nhập của người sử dụng. Kiểm soát nội dung thông tin lưu chuyển trên mạng.




**Hình 4 Một số chức năng của Firewall.**

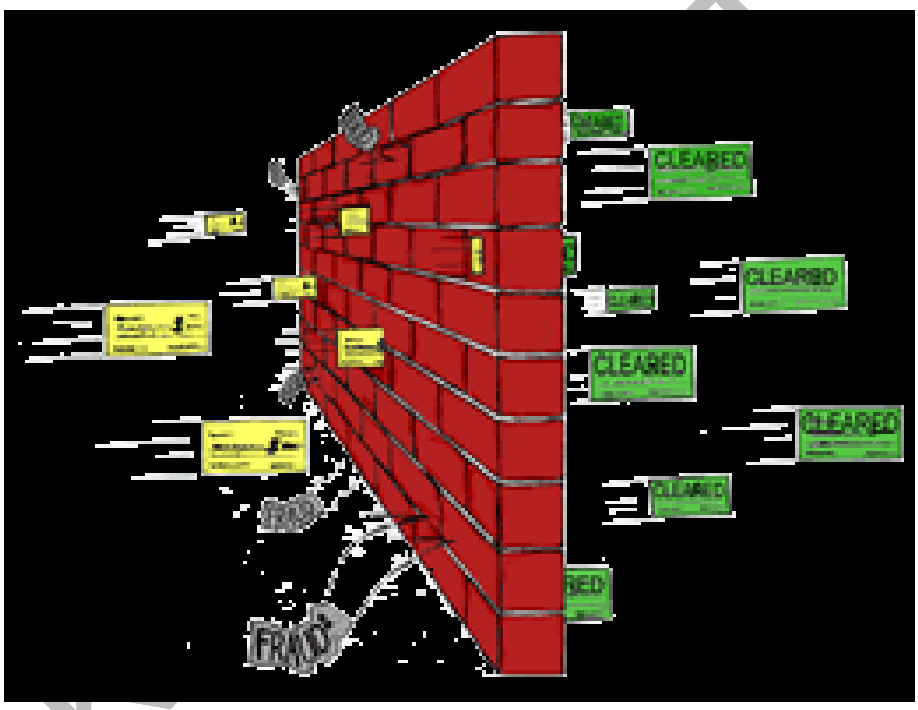
### 1.5.2 ***Thành phần:***

Firewall chuẩn bao gồm một hay nhiều các thành phần sau đây:

- Bộ lọc packet (packet-filtering router)
- Cổng ứng dụng (application-level gateway hay proxy server)
- Cổng mạch (circuite level gateway)
- Bộ lọc paket (Paket filtering router).

## 1.6 Nguyên lý:

Khi nói đến việc lưu thông dữ liệu giữa các mạng với nhau thông qua Firewall thì điều đó có nghĩa rằng Firewall hoạt động chặt chẽ với giao thức TCI/IP. Vì giao thức này làm việc theo thuật toán chia nhỏ các dữ liệu nhận được từ các ứng dụng trên mạng, hay nói chính xác hơn là các dịch vụ chạy trên các giao thức (Telnet, SMTP, DNS, SMNP, NFS...) thành các gói dữ liệu (data pakets) rồi gán cho các paket này những địa chỉ để có thể nhận dạng, tái lập lại ở đích cần gửi đến, do đó các loại Firewall cũng liên quan rất nhiều đến các packet và những con số địa chỉ của chúng.




**Hình 5 Lọc gói tin**

Bộ lọc packet cho phép hay từ chối mỗi packet mà nó nhận được. Nó kiểm tra toàn bộ đoạn dữ liệu để quyết định xem đoạn dữ liệu đó có thoả mãn một trong số các luật lệ của lọc packet hay không. Các luật lệ lọc packet này là dựa trên các thông tin ở đầu mỗi packet (packet header), dùng để cho phép truyền các packet đó ở trên mạng. Đó là:

- Địa chỉ IP nơi xuất phát ( IP Source address)
- Địa chỉ IP nơi nhận (IP Destination address)
- Những thủ tục truyền tin (TCP, UDP, ICMP, IP tunnel)
- Cổng TCP/UDP nơi xuất phát (TCP/UDP source port)
- Cổng TCP/UDP nơi nhận (TCP/UDP destination port)
- Dạng thông báo ICMP ( ICMP message type)
- Giao diện packet đến ( incomming interface of packet)
- Giao diện packet đi ( outcomming interface of packet)

Nếu luật lệ lọc packet được thoả mãn thì packet được chuyển qua Firewall. Nếu không packet sẽ bị bỏ đi. Nhờ vậy mà Firewall có thể ngăn cản được các kết nối vào các máy chủ hoặc mạng nào đó được xác định, hoặc khoá việc truy cập vào hệ thống mạng nội bộ từ những địa chỉ không cho phép. Hơn nữa, việc kiểm soát các cổng làm cho Firewall có khả năng chỉ cho phép một số loại kết nối nhất định vào các loại máy chủ nào đó, hoặc chỉ có những dịch vụ nào đó (Telnet, SMTP, FTP...) được phép mới chạy được trên hệ thống mạng cục bộ.

**Ưu điểm**:

- Đa số các hệ thống Firewall đều sử dụng bộ lọc packet. Một trong những ưu điểm của phương pháp dùng bộ lọc packet là chi phí thấp vì cơ chế lọc packet đã được bao gồm trong mỗi phần mềm router.
- Ngoài ra, bộ lọc packet là trong suốt đối với người sử dụng và các ứng dụng, vì vậy nó không yêu cầu sự huấn luyện đặc biệt nào cả.

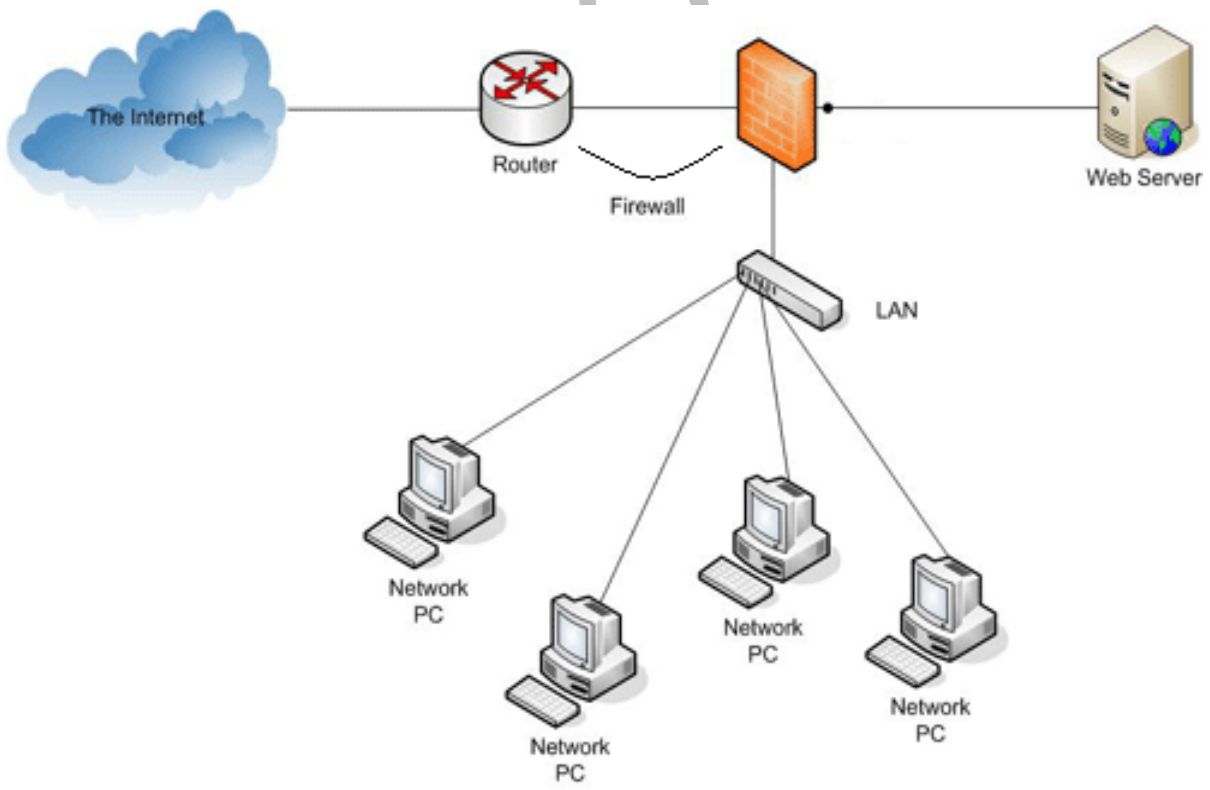
**Hạn chế:**

- Việc định nghĩa các chế độ lọc package là một việc khá phức tạp; đòi hỏi người quản trị mạng cần có hiểu biết chi tiết về các dịch vụ Internet, các dạng packet header, và các giá trị cụ thể có thể nhận trên mỗi trường. Khi

- Do làm việc dựa trên header của các packet, rõ ràng là bộ lọc packet không kiểm soát được nôi dung thông tin của packet. Các packet chuyển qua vẫn có thể mang theo những hành động với ý đồ ăn cắp thông tin hay phá hoại của kẻ xấu.

## 1.7 Các dạng firewall:

Mỗi dạng Firewall khác nhau có những thuận lợi và hạn chế riêng. Dạng phổ biến nhất là Firewall mức mạng (Network-level firewall). Loại Firewall này thường dựa trên bộ định tuyến, vì vậy các quy tắc quy định tính hợp pháp cho việc truy nhập được thiết lập ngay trên bộ định tuyến. Mô hình Firewall này sử dụng kỹ thuật lọc gói tin (packet-filtering technique), đó là tiến trình kiểm soát các gói tin qua bộ định tuyến.




**Hình 6 Firewall được cấu hình tại router**

Khi hoạt động, Firewall sẽ dựa trên bộ định tuyến mà kiểm tra địa chỉ nguồn (source address) hay địa chỉ xuất phát của gói tin. Sau khi nhận diện xong, mỗi địa chỉ nguồn IP sẽ được kiểm tra theo các quy tắc do người quản trị mạng định trước.

Firewall dựa trên bộ định tuyến làm việc rất nhanh do nó chỉ kiểm tra lướt trên các địa chỉ nguồn mà không hề có yêu cầu thực sự nào đối với bộ định tuyến, không tốn thời gian xử lý những địa chỉ sai hay không hợp lệ. Tuy nhiên, bạn phải trả giá: ngoại trừ những điều khiển chống truy nhập, các gói tin mang địa chỉ giả mạo vẫn có thể thâm nhập ở một mức nào đó trên máy chủ của bạn.

Một số kỹ thuật lọc gói tin có thể được sử dụng kết hợp với Firewall để khắc phục nhược điểm nói trên. Địa chỉ IP không phải là thành phần duy nhất của gói tin có thể mắc bẫy bộ định tuyến. Người quản trị nên áp dụng đồng thời các quy tắc, sử dụng thông tin định danh kèm theo gói tin như thời gian, giao thức, cổng... để tăng cường điều kiện lọc. Tuy nhiên, sự yếu kém trong kỹ thuật lọc gói tin của Firewall dựa trên bộ định tuyến không chỉ có vậy.

Một số dịch vụ gọi thủ tục từ xa (Remote Procedure Call - RPC) rất khó lọc một cách hiệu quả do các server liên kết phụ thuộc vào các cổng được gán ngẫu nhiên khi khởi động hệ thống. Dịch vụ gọi là ánh xạ cổng (portmapper) sẽ ánh xạ các lời gọi tới dịch vụ RPC thành số dịch vụ gán sẵn, tuy nhiên, do không có sự tương ứng giữa số dịch vụ với bộ định tuyến lọc gói tin, nên bộ định tuyến không nhận biết được dịch vụ nào dùng cổng nào, vì thế nó không thể ngăn chặn hoàn toàn các dịch vụ này, trừ khi bộ định tuyến ngăn toàn bộ các gói tin UDP (các dịch vụ RPC chủ yếu sử dụng giao thức UDP hay User Datagram Protocol). Việc ngăn chặn tất cả các gói tin UDP cũng sẽ ngăn luôn cả các dịch vụ cần thiết, ví dụ như DNS (Domain Name Service ố dịch vụ đặt tên vùng). Vì thế, dẫn đến tình trạng tiến thoái lưỡng nan.

## 1.8 Các ý niệm chung về Firewall:

Một trong những ý tưởng chính của Firewall là che chắn cho mạng của bạn khỏi tầm nhìn của những người dùng bên ngoài không được phép kết nối, hay chí ít cũng không cho phép họ rớ tới mạng. Quá trình này thực thi các chỉ tiêu lọc bỏ do người quản trị ấn định.

Trên lý thuyết, Firewall là phương pháp bảo mật an toàn nhất khi mạng của bạn có kết nối Internet. Tuy nhiên, vẫn tồn tại các vấn đề xung quanh môi trường bảo mật này. Nếu Firewall được cấu hình quá chặt chẽ, tiến trình làm việc của mạng sẽ bị ảnh hưởng, đặc biệt trong môi trường người dùng phụ thuộc hoàn toàn vào ứng dụng phân tán. Do Firewall thực thi từng chính sách bảo mật chặt chẽ nên nó có thể bị sa lầy. Tóm lại, cơ chế bảo mật càng chặt chẽ bao nhiêu, thì tính năng càng bị hạn chế bấy nhiêu.

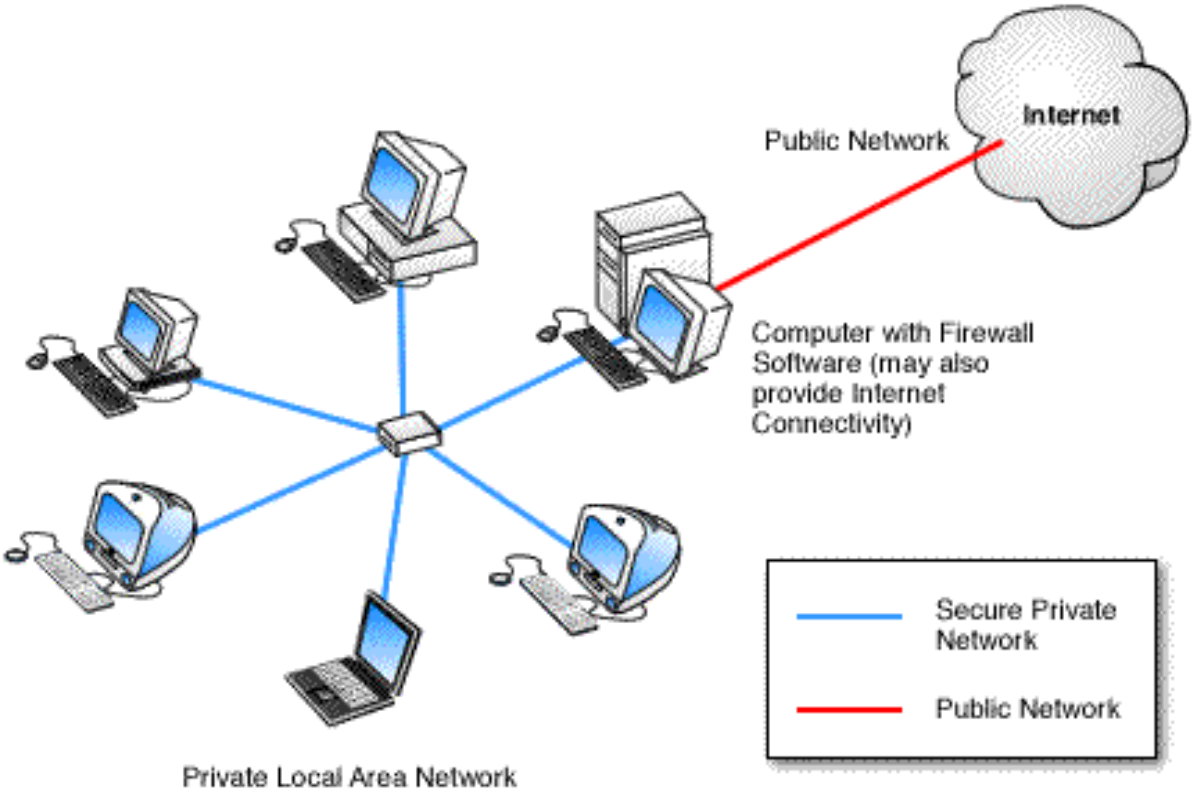
Một vấn đề khác của Firewall tương tự như việc xếp trứng vào rổ. Do là rào chắn chống kết nối bất hợp pháp nên một khe hở cũng có thể dễ dàng phá huỷ mạng của bạn. Firewall duy trì môi trường bảo mật, trong đó nó đóng vai trò điều khiển truy nhập và thực thi sơ đồ bảo mật. Firewall thường được mô tả như cửa ngõ của mạng, nơi xác nhận quyền truy nhập. Tuy nhiên điều gì sẽ xảy ra khi nó bị vô hiệu hoá? Nếu một kỹ thuật phá Firewall được phát hiện, cũng có nghĩa người vệ sĩ bị tiêu diệt và cơ hội sống sót của mạng là rất mỏng manh. Vì vậy trước khi xây dựng Firewall, bạn nên xem xét kỹ và tất nhiên phải hiểu tường tận về mạng của mình.

Một điều nữa, Firewall cũng có khả năng cấm các kết nối không được cho phép từ bên trong ra. Điều này, nếu suy nghĩ đơn giản thì chúng ta thấy rất có lợi, tuy nhiên trong một vài trường hợp thì nó vẫn có mặt hạn chế của nó.

### 1.8.1 ***Firewall dựa trên Application gateway:***

Một dạng phổ biến là Firewall dựa trên ứng dụng application-proxy. Loại này hoạt động hơi khác với Firewall dựa trên bộ định tuyến lọc gói tin. Application gateway dựa trên cơ sở phần mềm. Khi một người dùng không xác định kết nối từ

xa vào mạng chạy application gateway, gateway sẽ ngăn chặn kết nối từ xa này. Thay vì nối thông, gateway sẽ kiểm tra các thành phần của kết nối theo những quy tắc định trước. Nếu thoả mãn các quy tắc, gateway sẽ tạo cầu nối (bridge) giữa trạm nguồn và trạm đích.




**Hình 7 Firewall mềm**

Cầu nối đóng vai trò trung gian giữa hai giao thức. Ví dụ, trong một mô hình gateway đặc trưng, gói tin theo giao thức IP không được chuyển tiếp tới mạng cục bộ, lúc đó sẽ hình thành quá trình dịch mà gateway đóng vai trò bộ phiên dịch.

Ưu điểm của Firewall application gateway là không phải chuyển tiếp IP. Quan trọng hơn, các điều khiển thực hiện ngay trên kết nối. Sau cùng, mỗi công cụ đều cung cấp những tính năng thuận tiện cho việc truy nhập mạng. Do sự lưu chuyển của các gói tin đều được chấp nhận, xem xét, dịch và chuyển lại nên

Firewall loại này bị hạn chế về tốc độ. Quá trình chuyển tiếp IP diễn ra khi một server nhận được tín hiệu từ bên ngoài yêu cầu chuyển tiếp thông tin theo định dạng IP vào mạng nội bộ. Việc cho phép chuyển tiếp IP là lỗi không tránh khỏi, khi đó, hacker có thể thâm nhập vào trạm làm việc trên mạng của bạn.

Hạn chế khác của mô hình Firewall này là mỗi ứng dụng bảo mật (proxy application) phải được tạo ra cho từng dịch vụ mạng. Như vậy một ứng dụng dùng cho Telnet, ứng dụng khác dùng cho HTTP, v.v..

Do không thông qua quá trình chuyển dịch IP nên gói tin IP từ địa chỉ không xác định sẽ không thể tới máy tính trong mạng của bạn, do đó hệ thống application gateway có độ bảo mật cao hơn.

### 1.8.2 ***Cổng vòng(Circuit level gateway):***

Cổng vòng là một chức năng đặc biệt có thể thực hiện được bởi một cổng ứng dụng(application gateway). Cổng vòng đơn giản chỉ chuyển tiếp (relay) các kết nối TCP mà không thực hiện bất kỳ một hành động xử lý hay lọc packet nào.

VD: Cổng vòng đơn giản chuyển tiếp kết nối telnet qua firewall mà không thực hiện một sự kiểm tra, lọc hay điều khiển các thủ tục Telnet nào.Cổng vòng làm việc như một sợi dây, sao chép các byte giữa kết nối bên trong (inside connection) và các kết nối bên ngoài (outside connection). Tuy nhiên, vì sự kết nối này xuất hiện từ hệ thống firewall, nên nó che dấu thông tin về mạng nội bộ.

Cổng vòng thường được sử dụng cho những kết nối ra ngoài, nơi mà các quản trị mạng thật sự tin tưởng những người dùng bên trong. Ưu điểm lớn nhất là một bastion host có thể được cấu hình như là một hỗn hợp cung cấp Cổng ứng dụng cho những kết nối đến, và cổng vòng cho các kết nối đi. Điều này làm cho hệ thống Firewall dễ dàng sử dụng cho những người trong mạng nội bộ muốn trực tiếp truy nhập tới các dịch vụ Internet, trong khi vẫn cung cấp chức năng Firewall để bảo vệ mạng nội bộ từ những sự tấn công bên ngoài.

### 1.8.3 ***Hạn chế của Firewall:***

- Firewall không đủ thông minh như con ngời để có thể đọc hiểu từng loại thông tin và phân tích nội dung tốt hay xấu của nó. Firewall chỉ có thể ngăn chặn sự xâm nhập của những nguồn thông tin không mong muốn nhưng phải xác định rõ các thông số địa chỉ.
- Firewall không thể ngăn chặn một cuộc tấn công nếu cuộc tấn công này không "đi qua" nó. Một cách cụ thể, firewall không thể chống lại một cuộc tấn công từ một đường dial-up, hoặc sự dò rỉ thông tin do dữ liệu bị sao chép bất hợp pháp lên đĩa mềm.
- Firewall cũng không thể chống lại các cuộc tấn công bằng dữ liệu (data-driven attack). Khi có một số chương trình được chuyển theo thư điện tử, vượt qua firewall vào trong mạng được bảo vệ và bắt đầu hoạt động ở đây.
- Một ví dụ là các virus máy tính. Firewall không thể làm nhiệm vụ rà quét virus trên các dữ liệu được chuyển qua nó, do tốc độ làm việc, sự xuất hiện liên tục của các virus mới và do có rất nhiều cách để mã hóa dữ liệu, thoát khỏi khả năng kiểm soát của firewall.

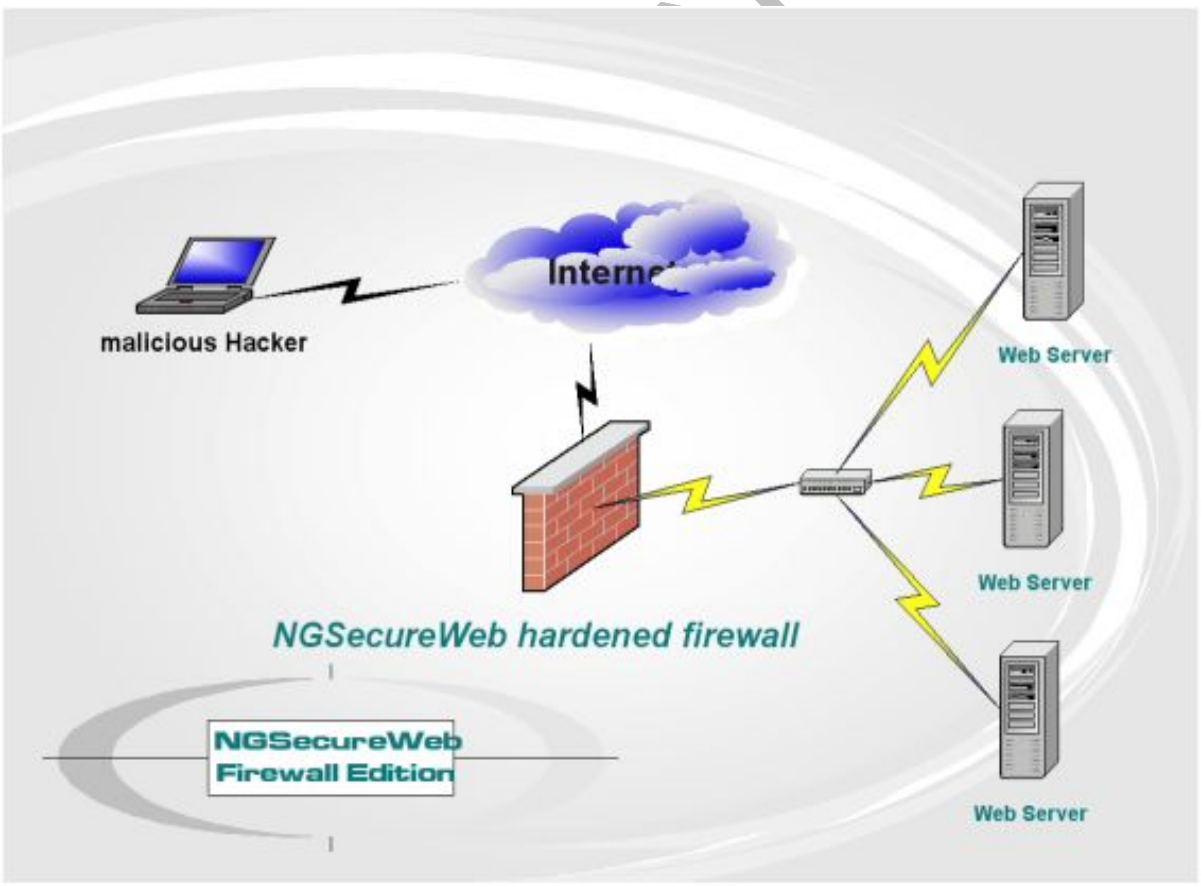
  Tuy nhiên, Firewall vẫn là giải pháp hữu hiệu được áp dụng rộng rãi.

### 1.8.4 ***Firewall có dễ phá hay không:***

Câu trả lời là không. Lý thuyết không chứng minh được có khe hở trên Firewall, tuy nhiên thực tiễn thì lại có. Các hacker đã nghiên cứu nhiều cách phá Firewall. Quá trình phá Firewall gồm hai giai đoạn: đầu tiên phải tìm ra dạng Firewall mà mạng sử dụng cùng các loại dịch vụ hoạt động phía sau nó; tiếp theo là phát hiện khe hở trên Firewall đó, giai đoạn này thường khó khăn hơn. Theo nghiên cứu của các hacker, khe hở trên Firewall tồn tại là do lỗi định cấu hình của người quản trị hệ thống, sai sót này cũng không hiếm khi xảy ra. Người quản trị phải chắc

chắn sẽ không có bất trắc cho dù sử dụng hệ điều hành (HĐH) mạng nào, đây là cả một vấn đề nan giải. Trong các mạng UNIX, điều này một phần là do HĐH UNIX quá phức tạp, có tới hàng trăm ứng dụng, giao thức và lệnh riêng. Sai sót trong xây dựng Firewall có thể do người quản trị mạng không nắm vững về TCP/IP.

Một trong những việc phải làm của các hacker là tách các thành phần thực ra khỏi các thành phần giả mạo. Nhiều Firewall sử dụng trạm hy sinh (sacrificial hosts) - là hệ thống được thiết kế như các server Web (có thể sẵn sàng bỏ đi) hay bẫy (decoys), dùng để bắt các hành vi thâm nhập của hacker. Bẫy có thể cần dùng tới những thiết bị ngụy trang phức tạp nhằm che dấu tính chất thật của nó, ví dụ: đưa ra câu trả lời tương tự hệ thống tập tin hay các ứng dụng thực. Vì vậy, công việc đầu tiên của hacker là phải xác định đây là các đối tượng tồn tại thật.




**Hình 8 Tấn công hệ thống từ bên ngoài**

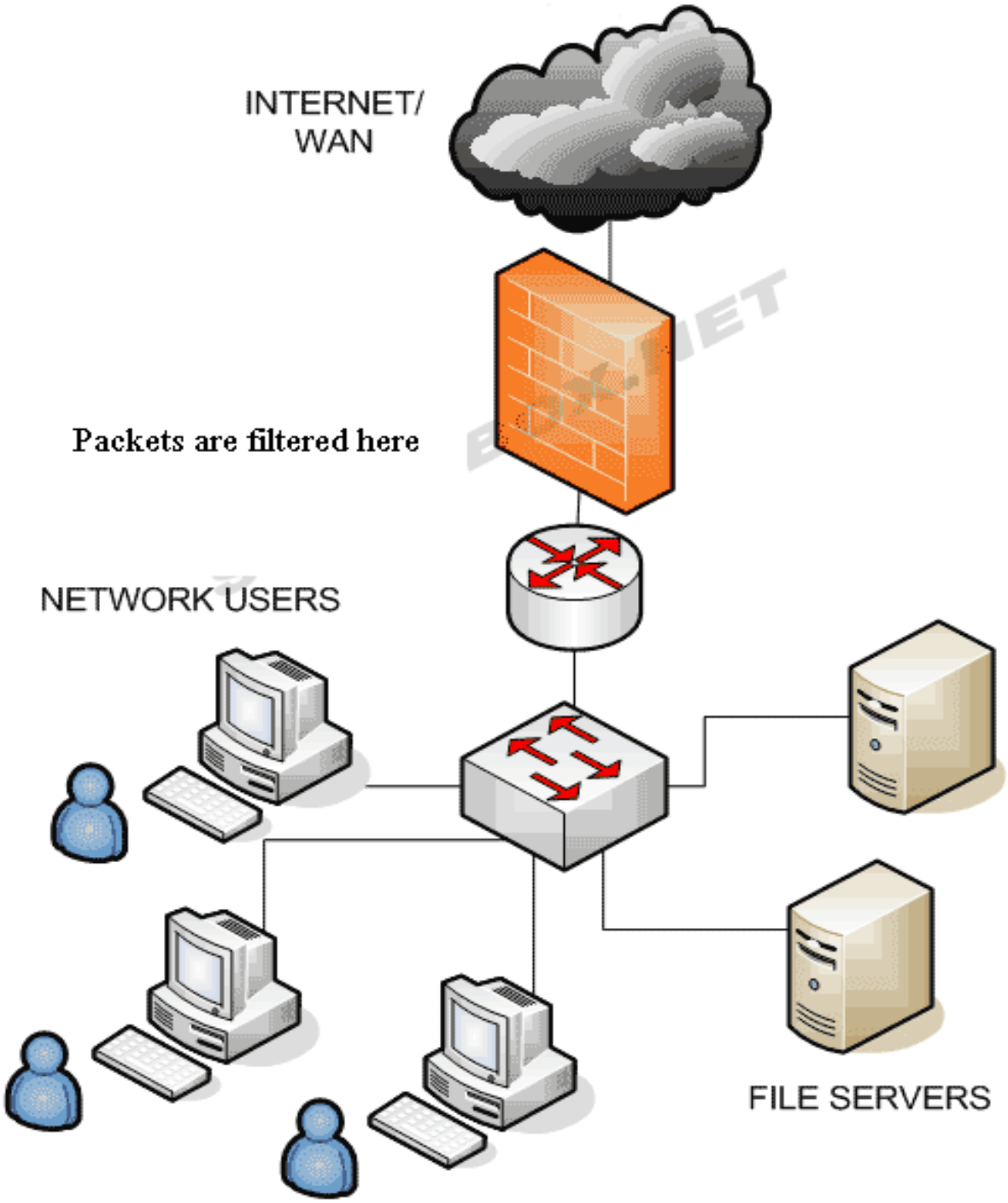
Để có được thông tin về hệ thống, hacker cần dùng tới thiết bị có khả năng phục vụ mail và các dịch vụ khác. Hacker sẽ tìm cách để nhận được một thông điệp đến từ bên trong hệ thống, khi đó, đường đi được kiểm tra và có thể tìm ra những manh mối về cấu trúc hệ thống.

Ngoài ra, không Firewall nào có thể ngăn cản việc phá hoại từ bên trong. Nếu hacker tồn tại ngay trong nội bộ tổ chức, chẳng bao lâu mạng của bạn sẽ bị hack. Thực tế đã xảy ra với một công ty dầu lửa lớn: một tay hacker trà trộn vào đội ngũ nhân viên và thu thập những thông tin quan trọng không chỉ về mạng mà còn về các trạm Firewall.

## 1.9 Một số mô hình Firewall:

### 1.9.1 ***Packet-Filtering Router:***

Hệ thống Internet firewall phổ biến nhất chỉ bao gồm một packet-filtering router đặt giữa mạng nội bộ và Internet. Một packet-filtering router có hai chức năng: chuyển tiếp truyền thông giữa hai mạng và sử dụng các quy luật về lọc gói để cho phép hay từ chối truyền thông.




**Hình 9 Packet filtering**

Căn bản, các quy luật lọc được định nghĩa sao cho các host trên mạng nội bộ được quyền truy nhập trực tiếp tới Internet, trong khi các host trên Internet chỉ có một số giới hạn các truy nhập vào các máy tính trên mạng nội bộ. Tư tưởng của mô cấu trúc firewall này là tất cả những gì không được chỉ ra rõ ràng là cho phép thì có nghĩa là bị từ chối.

**Ưu điểm**:

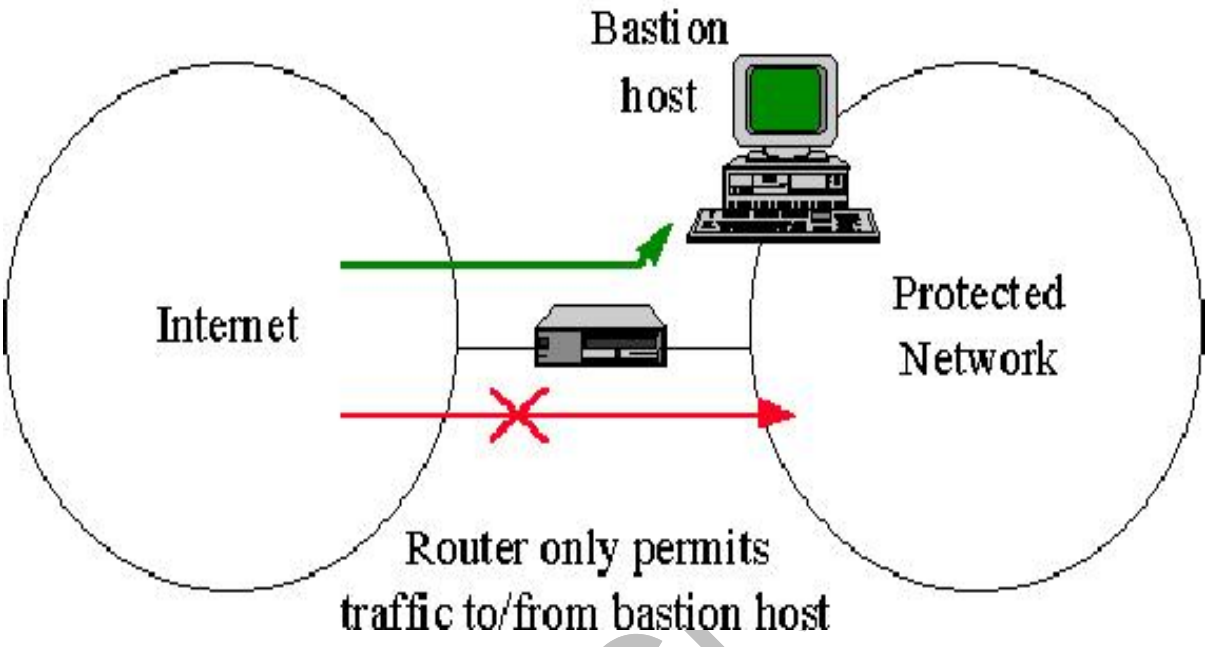
- Giá thành thấp, cấu hình đơn giản
- Trong suốt(transparent) đối với user.

**Hạn chế**:

- Có rất nhiều hạn chế đối với một packet-filtering router, như là dễ bị tấn công vào các bộ lọc mà cấu hình được đặt không hoàn hảo, hoặc là bị tấn công ngầm dưới những dịch vụ đã được phép.
- Bởi vì các packet được trao đổi trực tiếp giữa hai mạng thông qua router, nguy cơ bị tấn công quyết định bởi số lợng các host và dịch vụ được phép. Điều đó dẫn đến mỗi một host được phép truy nhập trực tiếp vào Internet cần phải được cung cấp một hệ thống xác thực phức tạp, và thường xuyên kiểm tra bởi người quản trị mạng xem có dấu hiệu của sự tấn công nào không.
- Nếu một packet-filtering router do một sự cố nào đó ngừng hoạt động, tất cả hệ thống trên mạng nội bộ có thể bị tấn công.

### 1.9.2 ***Mô hình Single-Homed Bastion Host:***

Hệ thống này bao gồm một packet-filtering router và một bastion host. Hệ thống này cung cấp độ bảo mật cao hơn hệ thống trên, vì nó thực hiện cả bảo mật ở tầng network (packet-filtering) và ở tầng ứng dụng (application level). Đồng thời, kẻ tấn công phải phá vỡ cả hai tầng bảo mật để tấn công vào mạng nội bộ.




**Hình 10 Mô hình single-Homed Bastion Host**

Trong hệ thống này, bastion host được cấu hình ở trong mạng nội bộ. Qui luật filtering trên packet-filtering router được định nghĩa sao cho tất cả các hệ thống ở bên ngoài chỉ có thể truy nhập bastion host; Việc truyền thông tới tất cả các hệ thống bên trong đều bị khoá. Bởi vì các hệ thống nội bộ và bastion host ở trên cùng một mạng, chính sách bảo mật của một tổ chức sẽ quyết định xem các hệ thống nội bộ được phép truy nhập trực tiếp vào bastion Internet hay là chúng phải sử dụng dịch vụ proxy trên bastion host. Việc bắt buộc những user nội bộ được thực hiện bằng cách đặt cấu hình bộ lọc của router sao cho chỉ chấp nhận những truyền thông nội bộ xuất phát từ bastion host.

**Ưu điểm:**

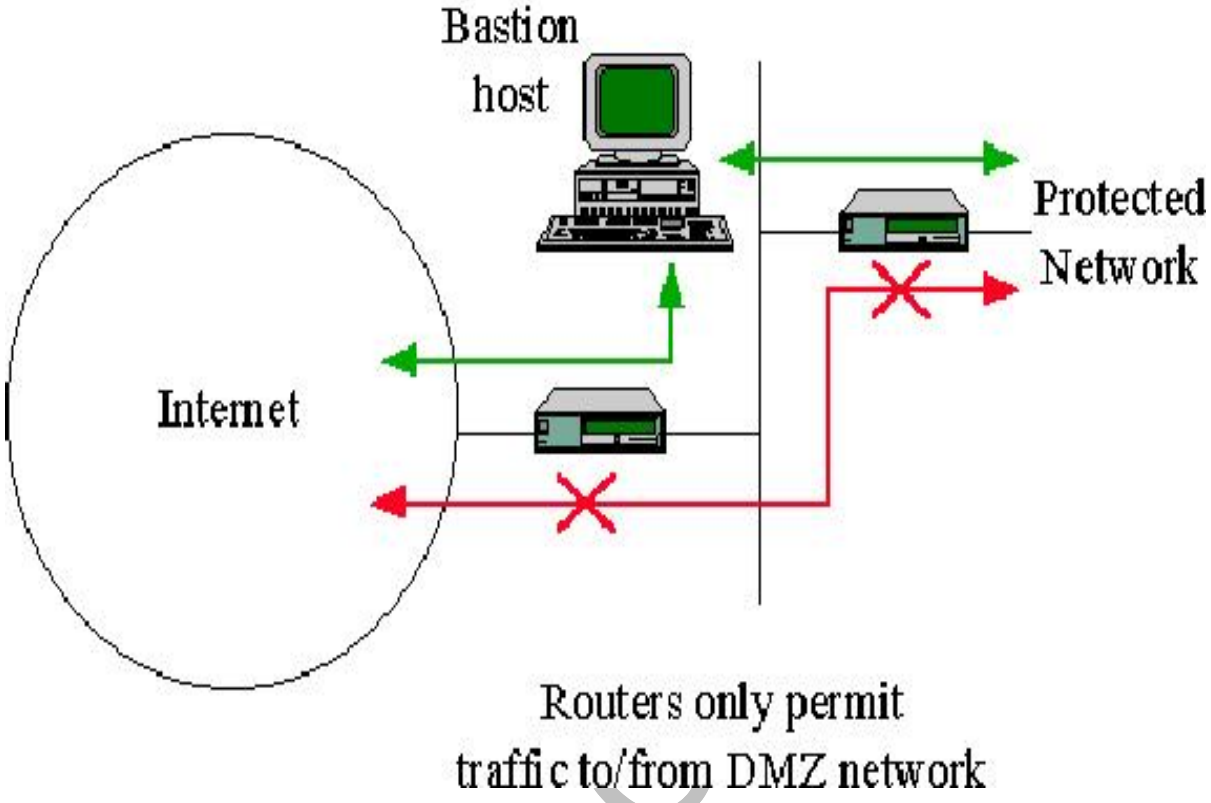
Máy chủ cung cấp các thông tin công cộng qua dịch vụ Web và FTP có thể đặt trên packet-filtering router và bastion. Trong trường hợp yêu cầu độ an toàn cao nhất, bastion host có thể chạy các dịch vụ proxy yêu cầu tất cả các

Bởi vì bastion host là hệ thống bên trong duy nhất có thể truy nhập được từ Internet, sự tấn công cũng chỉ giới hạn đến bastion host mà thôi. Tuy nhiên, nếu như user log on được vào bastion host thì họ có thể dễ dàng truy nhập toàn bộ mạng nội bộ. Vì vậy cần phải cấm không cho user logon vào bastion host.

### 1.9.3 ***Mô hình Dual-Homed Bastion Host:***

Demilitarized Zone (DMZ) hay Screened-subnet Firewall

Hệ thống bao gồm hai packet-filtering router và một bastion host. Hệ có độ an toàn cao nhất vì nó cung cấp cả mức bảo mật network và application, trong khi định nghĩa một mạng "phi quân sự". Mạng DMZ đóng vai trò như một mạng nhỏ, cô lập đặt giữa Internet và mạng nội bộ. Cơ bản, một DMZ được cấu hình sao cho các hệ thống trên Internet và mạng nội bộ chỉ có thể truy nhập được một số giới hạn các hệ thống trên mạng DMZ, và sự truyền trực tiếp qua mạng DMZ là không thể được.




**Hình 11 Mô hình Dual-Homed Bastion Host**

Với những thông tin đến, router ngoài chống lại những sự tấn công chuẩn (như giả mạo địa chỉ IP), và điều khiển truy nhập tới DMZ. Hệ thống chỉ cho phép bên ngoài truy nhập vào bastion host. Router trong cung cấp sự bảo vệ thứ hai bằng cách điều khiển DMZ truy nhập mạng nội bộ chỉ với những truyền thông bắt đầu từ bastion host.

Với những thông tin đi, router trong điều khiển mạng nội bộ truy nhập tới DMZ. Nó chỉ cho phép các hệ thống bên trong truy nhập bastion host và có thể cả information server. Quy luật filtering trên router ngoài yêu cầu sử dung dich vụ proxy bằng cách chỉ cho phép thông tin ra bắt nguồn từ bastion host.

**Ưu điểm:**

Kẻ tấn công cần phá vỡ ba tầng bảo vệ: router ngoài, bastion host và router

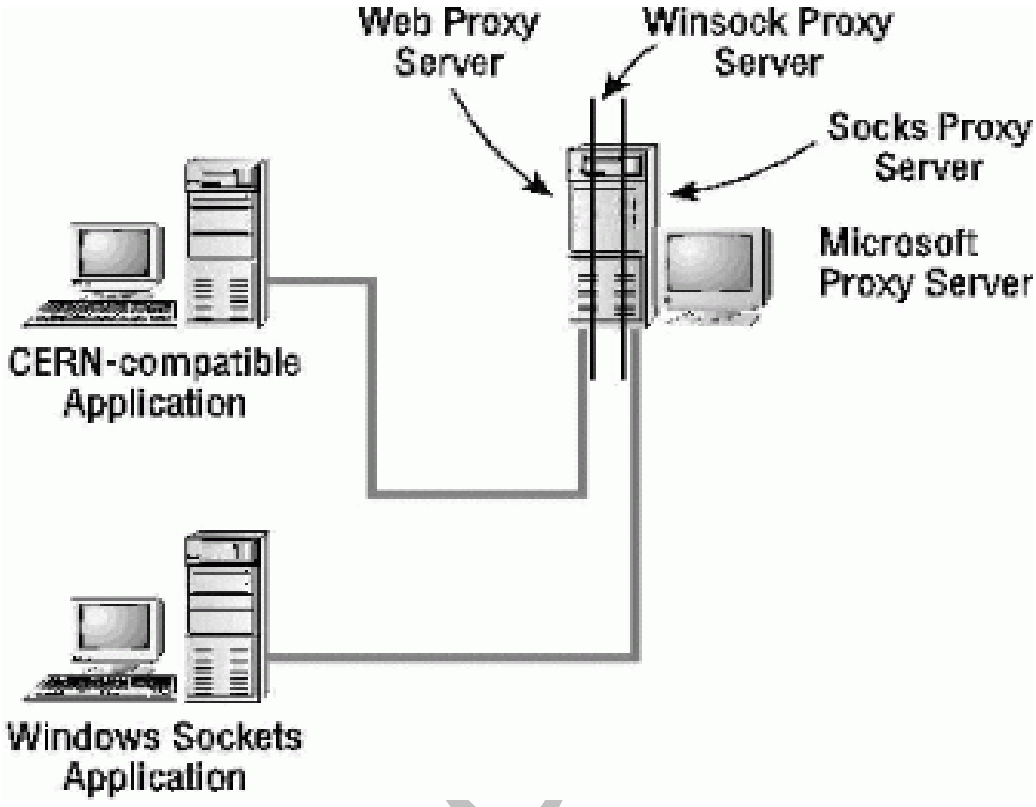
Chỉ có một số hệ thống đã được chọn ra trên DMZ là được biết đến bởi Internet qua routing table và DNS information exchange ( Domain Name Server ).

Bởi vì router trong chỉ quảng cáo DMZ network tới mạng nội bộ, các hệ thống trong mạng nội bộ không thể truy nhập trực tiếp vào Internet. Điều nay đảm bảo rằng những user bên trong bắt buộc phải truy nhập Internet qua dịch vụ proxy.

### 1.9.4 ***Proxy server:***

Chúng ta sẽ xây dựng Firewall theo kiến trúc application-level gateway, theo đó một bộ chương trình proxy được đặt ở gateway ngăn cách một mạng bên trong (Intranet) với Internet.

Bộ chương trình proxy được phát triển dựa trên bộ công cụ xây dựng Internet Firewall TIS (Trusted Information System), bao gồm một bộ các chương trình và sự đặt lại cấu hình hệ thống để nhằm mục đích xây dựng một Firewall. Bộ chương trình được thiết kế để chạy trên hệ UNIX sử dụng TCP/IP với giao diện socket Berkeley.




**Hình 12 Mô hình 1 Proxy đơn giản**

Bộ chương trình proxy được thiết kế cho một số cấu hình firewall, theo các dạng cơ bản: dual-home gateway, screened host gateway, và screened subnet gateway.

Thành phần Bastion host trong Firewall, đóng vai trò như một người chuyển tiếp thông tin, ghi nhật ký truyền thông, và cung cấp các dịch vụ, đòi hỏi độ an toàn cao.

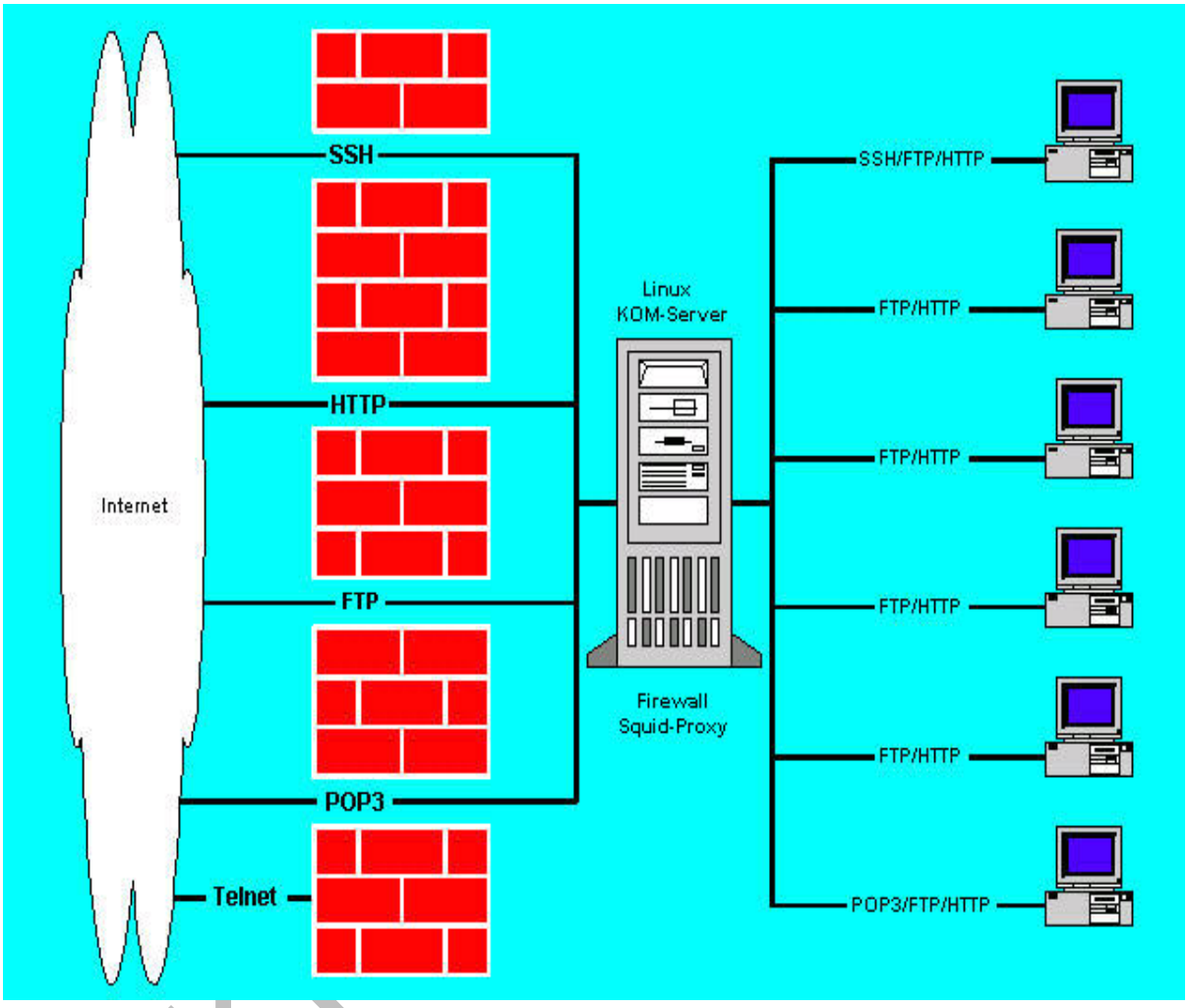
Proxy server chúng ta sẽ tìm hiểu kĩ hơn ở phần sau.

### 1.9.5 ***Phần mềm Firewall – Proxy server:***

Bộ chương trình proxy gồm những chương trình mức ứng dụng (application-level programs), dùng để thay thế hoặc là thêm vào phần mềm hệ thống. Đối với mỗi dịch vụ, cần có một phần mềm tương ứng làm nhiệm vụ lọc các bản tin. Trên

- SMTP Gateway - Proxy server cho dịch vụ SMTP (Simple Mail Tranfer Protocol)
- FTP Gateway - Proxy server cho dịch vụ Ftp
- Telnet Gateway - Proxy server cho dịch vụ Telnet
- HTTP Gateway - Proxy server cho dịch vụ HTTP (World Wide Web)
- Rlogin Gateway - Proxy server cho dịch vu rlogin
- Plug Gateway - Proxy server cho dịch vụ kết nối server tức thời dùng giao thức TCP (TCP Plug-Board Connection server)
- SOCKS - Proxy server cho các dịch vụ theo chuẩn SOCKS
- NETACL - Điều khiển truy nhập mạng dùng cho các dịch vụ khác
- IP filter – Proxy điều khiển mức IP
- SMTP Gateway - Proxy server cho cổng SMTP

#### 1.9.5.1 SMTP Gateway - Proxy server cho dịch vụ SMTP (Simple Mail Tranfer Protocol)




**Hình 13 Một số protocol sau proxy**

Chương trình SMTP Gateway được xây dựng trên cơ sở sử dụng hai phần mềm smap và smapd, dùng để chống lại sự truy nhập thông qua giao thức SMTP. Nguyên lý thực hiện là chặn trước chương trình mail server nguyên thuỷ của hệ thống, không cho phép các hệ thống bên ngoài kết nối trực tiếp với mail server. Vì ở trong mạng tin cậy mail server thường có một số quyền

Khi một hệ thống ở xa nối tới cổng SMTP. Chương trình smap sẽ dành quyền phục vụ và chuyển tới thư mục dành riêng và đặt user-id ở mức bình thường (không có quyền ưu tiên). Mục đích duy nhất của smap là đối thoại SMTP với các hệ thống khác, thu lượm mail, ghi vào đĩa, ghi nhật ký, và kết thúc. Smapd thường xuyên quét thư mục này, khi phát hiện có thư sẽ chuyển dữ liệu cho sendmail để phân phát vào các hòm thư cá nhân hoặc chuyển tiếp tới các mail server khác.

Như vậy, một user lạ trên mạng không thể kết nối trực tiếp với Mail Server. Tất cả các thông tin đi theo đường này hoàn toàn có thể kiểm soát được. Tuy nhiên, chương trình cũng không thể giải quyết vấn đề giả mạo thư hoặc các loại tấn công bằng đường khác.

#### 1.9.5.2 FTP Gateway Proxy Server cho dịch vụ FTP:

Proxy server cho dịch vụ FTP cung cấp khả năng kiểm soát truy nhập dịch vụ FTP dựa trên địa chỉ IP và hostname, và cung cấp điều khiển truy nhập thứ cấp cho phép tuỳ chọn khoá hoặc ghi nhật ký bất kỳ lệnh FTP nào. Các địa chỉ đích của dịch vụ này cũng có thể tuỳ chọn được phép hay bị cấm. Tất cả các sự kết nối và dung lượng dữ liệu chuyển qua đều bị ghi nhật kí lại.

FTP Gateway tự bản thân nó không đe dọa an toàn của hệ thống Firewall, bởi vì nó chạy tới một thư mục rỗng và không thực hiện một thủ tục vào ra file nào cả ngoài việc đọc file cấu hình của nó.

FTP Server chỉ cung cấp dịch vụ FTP, mà không quan tâm đến ai có quyền hay không có quyền kết xuất (download) file. Do vậy, việc xác định quyền phải được thiết lập trên FTP Gateway và phải thực hiện trước khi thực hiện việc kết xuất (download) hay nhập (upload) file. Ftp Gateway nên được cấu

#### 1.9.5.3 Telnet Gateway Proxy Server cho dịch vụ Telnet:

Telnet Gateway là một proxy server quản lý truy nhập mạng dựa trên địa chỉ IP và/hoặc hostname, và cung cấp sự điều khiển truy nhập thứ cấp cho phép tuỳ chọn khoá bất kỳ đích nào. Tất cả các sự kết nối dữ liệu chuyển qua đều được ghi nhật ký lại. Mỗi một lần user nối tới Telnet Gateway, người sử dụng phải lựa chọn phương thức kết nối.

Telnet Gateway không phương hại tới an toàn hệ thống, vì nó chỉ hoạt động trong một phạm vi cho phép nhất định. Cụ thể, hệ thống sẽ chuyển điều khiển tới một thư mục dành riêng. Đồng thời cấm truy nhập tới các thư mục và file khác.

Telnet Gateway được sử dụng để kiểm soát các truy nhập vào hệ thống mạng nội bộ. Các truy nhập không được phép sẽ không thể thực hiện được còn các truy nhập hợp pháp sẽ bị ghi lại nhật ký về thời gian truy nhập và các thao tác đã thực hiện.

**HTTP Gateway - Proxy server cho web**:

HTTP Gateway là một Proxy Server quản lý truy nhập hệ thống qua cổng HTTP (Web). Chơng trình này, dựa trên địa chỉ đích và địa chỉ nguồn để ngăn cấm hoặc cho phép yêu cầu truy nhập đi qua.

Đồng thời căn cứ và mã lệnh của giao thức HTTP, phần mềm này sẽ cho

**Rlogin Gateway - Proxy server cho rlogin**:

Các terminal truy nhập qua thủ tục BSD rlogin được kiểm soát bởi rlogin gateway. Chương trình cho phép kiểm tra và điều khiển truy nhập mạng tương tự như telnet gateway. Rlogin client có thể chỉ ra một hệ thống ở xa ngay khi bắt đầu nối vào proxy. Chương trình sẽ hạn chế yêu cầu tương tác giữa user với máy.

**Plug Gateway - TCP Plug-Board Connection server**:

Firewall cung cấp các dịch vụ thông thường như Usernet news. Người quản trị mạng có thể chọn hoặc là chạy dịch vụ này ngay trong firewall, hoặc cài đặt một proxy server cho dịch vụ này.

Do dịch vụ News chạy trực tiếp trên firewall thì dễ gây lỗi hệ thống, nên cách an toàn hơn là sử dụng proxy. Plug gateway được thiết kế để kiểm soát dịch vụ Usernet News và một số dịch vụ khác như Lotus Notes, Oracle, etc. Plug gateway dựa trên địa chỉ IP hoặc hostname, sẽ cho phép kiểm soát tất cả các truy nhập hệ thống thông qua các cổng dịch vụ được đăng ký. Trên cơ sở đó sẽ cho phép hoặc cấm các yêu cầu truy nhập. Tất cả yêu cầu kết nối bao gồm cả dữ liệu có thể được ghi lại nhật ký để theo dõi và kiểm soát.

#### 1.9.5.4 SQL Gateway Proxy Server cho SQL-Net:

SQL Net sử dụng giao thức riêng không giống như của News hay Lotus Notes, Do vậy, không thể sử dụng Plug Gateway cho dịch vụ này được. SQL

#### 1.9.5.5 SOCKS Gateway và NETACL:

**SOCKS Gateway - Proxy server cho các dịch vụ theo chuẩn SOCKS:**

SOCKS là giao thức kết nối mạng giữa các máy chủ cùng hỗ trợ giao thức này. Hai máy chủ khi sử dụng giao thức này sẽ không cần quan tâm tới việc giữa chúng có thể nối ghép thông qua IP hay không.

SOCKS sẽ địch hướng lại các yêu cầu ghép nối từ máy chủ đầu kia. Máy chủ SOCKS sẽ xác định quyền truy nhập và thiết lập kênh truyền thông tin giữa hai máy. SOCKS Gateway dùng để chống lại các truy nhập vào mạng thông qua cổng này.

**NETACL - Công cụ điều khiển truy nhập mạng**:

Các dịch vụ thông thường trên mạng không cung cấp khả năng kiểm soát truy cập tới chúng do vậy chúng là các điểm yếu để tấn công. Kể cả trên hệ thống firewall các dịch vụ thông thường đã được lọc bỏ khá nhiều để đảm bảo an toàn hệ thống nhưng một số dich vụ vẫn cần thiết để duy trì hệ thống như telnet, rlogin...

Netacl là một công cụ để điều khiển truy nhập mạng, dựa trên địa chỉ network của máy client, và dịch vụ đợc yêu cầu. Nó bao trùm nên các dịch vụ cơ bản cung cấp thêm khả năng kiểm soát cho dịch vụ đó. Vì vậy một client (xác định bởi địa chỉ IP hoặc hostname) có thể truy nhập tới telnet server khi nó nối với cổng dịch vụ telnet trên firewall.

Thường thường trong các cấu hình firewall, NETACL được sử dụng để cấm tất cả các máy trừ một vài host được quyền login tới firewall qua hoặc là telnet hoặc là rlogin, và để khoá các truy nhập từ những kẻ tấn công.
Độ an toàn của Netacl dựa trên địa chỉ IP và/hoặc hostname. Với các hệ thống cần độ an toàn cao, nên dụng địa chỉ IP để tránh sự giả mạo DNS.
Netacl không chống lại được sự giả địa chỉ IP qua chuyển nguồn (source routing) hoặc những phương tiện khác. Nếu có các loại tấn công như vậy, cần phải sử dụng một router có khả năng soi những packet đã được chuyển nguồn (screening source routed packages).
Chú ý là netacl không cung cấp điều khiển truy nhập UDP, bởi vì công nghệ hiện nay không đảm bảo sự xác thực của UDP. An toàn cho các dịch vụ UDP ở đây đồng nghĩa với sự không cho phép tất cả các dịch vụ UDP.

#### 1.9.5.6 Authentication:

Bộ Firewall chứa chương trình server xác thực được thiết kế để hỗ trợ cơ chế phân quyền. Authsrv chứa một cơ sở dữ liệu về người dùng trong mạng, mỗi bản ghi tương ứng với một ngời dùng, chứa cơ chế xác thực cho mỗi anh ta, trong đó bao gồm tên nhóm, tên đầy đủ của ngời dùng, lần truy cập mới nhất. Mật khẩu không mã hoá (Plain text password) được sử dụng cho người dùng trong mạng để việc quản trị được đơn giản. Mật khẩu không mã hoá không nên dùng với những ngòi sử dụng từ mạng bên ngoài.
Người dùng trong cơ sở dữ liệu của có thể được chia thành các nhóm khác nhau được quản trị bởi quản trị nhóm là người có toàn quyền trong nhóm cả việc thêm, bớt ngời dùng. Điều này thuận lợi khi nhiều tổ chức cùng dùng chung một Firewall.
Authsrv quản lý nhóm rất mềm dẻo, quản trị có thể nhóm người dùng thành nhóm dùng "group wiz", người có quyền quản trị nhóm có thể xoá, thêm, tạo

#### 1.9.5.7 IP Filter – Bộ lọc mức IP:

IP Filter là bộ lọc các gói tin TCP/IP, được xem như thành phần không thể thiếu khi thiết lập Firewall trong suốt đối với ngời sử dụng. Phần mềm này sẽ đợc cài đặt trong lõi của hệ thống (như UNIX kernel), được chạy ngầm khi hệ thống hoạt động, để đón nhận và phân tích tất cả các gói IP (IP Package).

**Bộ lọc IP filter có thể thực hiện các việc sau:**

- Cho đi qua hoặc cấm bất kỳ một gói tin nào.
- Nhận biết được các dịch vụ khác nhau
- Lọc theo địa chỉ IP hoặc hosts
- Cho phép lọc chọn lựa giao thức IP bất kỳ
- Cho phép lọc chọn lựa theo các mảnh IP
- Cho phép lọc chọn lựa theo các tuỳ chọn IP
- Gửi trả lại các khối ICMP/TCP lỗi và đặt lại số hiệu packet
- Lưu giữ các thông tin trạng thái đối với các dòng TCP, UDP and ICMP
- Lưu giữ các thông tin trạng thái đối với các mảnh IP packet bất kỳ
- Có chức năng như Network Address Translator (NAT)

- Làm cơ sở thiết lập các kết nối trong suốt đối với người sử dụng
- Cung cấp các header cho các chương trình của người sử dụng để xác nhận.
- Ngoài ra hỗ trợ không gian tạm cho các quy tắc xác nhận đối với các gói tin đi qua.

  Đặc biệt đối với các giao thức cơ bản của Internet, TCP, UDP và ICMP, thì IP filter cho phép lọc theo:

  - Inverted host/net matching
  - Số hiệu cổng của các gói tin TCP/UDP
  - Kiểu hoặc mã của các gói tin ICMP
  - Thiết lập các gói tin TCP
  - Tổ hợp tuỳ ý các cờ trạng thái TCP
  - Lọc/loại bỏ những gói IP cha kết thúc
  - Lọc theo kiểu dịch vụ
  - Cho phép ghi nhật ký các bản tin bao gồm:
- Header của các gói tin TCP/UDP/ICMP and IP
- Một phần hoặc tất cả dữ liệu của gói tin

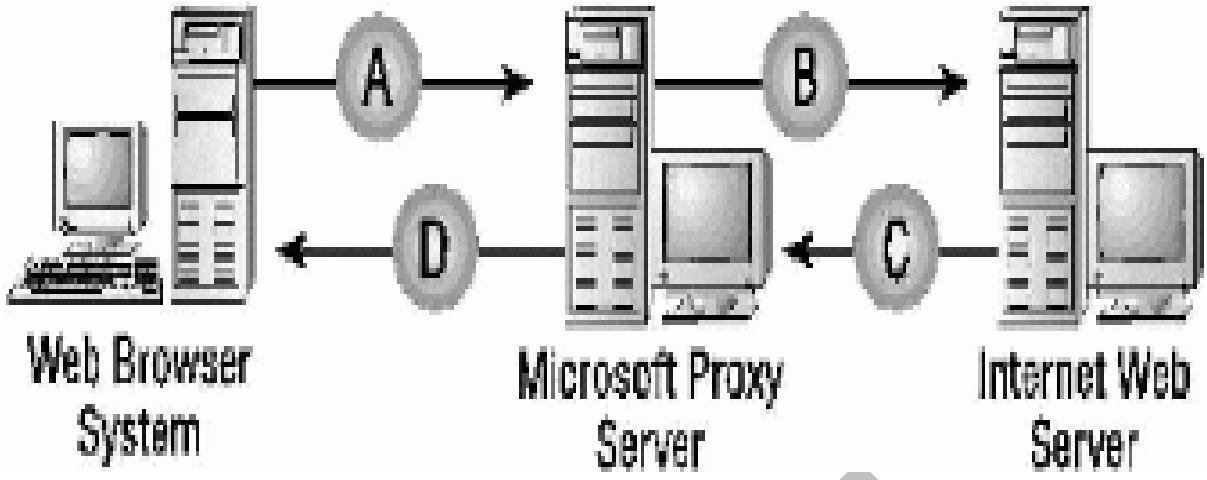
## 1.10 Lời kết:

Hiện tại, Firewall là phương pháp bảo vệ mạng phổ biến nhất, 95% cộng đồng hacker phải thừa nhận là dường như không thể vượt qua Firewall. Song trên thực tế, Firewall đã từng bị phá. Nếu mạng của bạn có kết nối Internet và chứa dữ liệu quan trọng cần được bảo vệ, bên cạnh Firewall, bạn nên tăng cường các biện pháp bảo vệ khác như là bảo mật ở mức physical, thường xuyên back up dữ liệu, chọn lọc nhân viên…

# Chương 2: KHÁI NIỆM PROXY

## 2.1 Proxy là gì:

- Theo www.learnthat.com: proxy là một thiết bị cho phép kết nối vào internet, nó đứng giữa các workstation trong một mạng và internet, cho phép bảo mật kết nối, chỉ cho phép một số cổng và protocol nào đó, vd: tcp, http, telnet trên các cổng 80, 23…. Khi một client yêu cầu một trang nào đó, yêu cầu này sẽ được chuyển đến proxy server, proxy server sẽ chuyển tiếp yêu cầu này đến site đó. Khi yêu cầu được đáp trả, proxy sẽ trả kết quả này lại cho client tương ứng. Proxy server có thể được dùng để ghi nhận việc sử dụng internet và ngăn chặn những trang bị cấm
- Theo www.nyu.edu: proxy server là một server đứng giữa một ứng dụng của client, như web browser, và một server ở xa (remote server). Proxy server xem xét các request xem nó có thể xử lý bằng cache của nó không, nếu không thể, nó sẽ chuyển yêu cầu này đến remote server.
- Theo www.webopedia.com: proxy server là một server đứng giữa một ứng dụng client, như web browser, và một server thực. Nó chặn tất cả các yêu cầu đến các server thực để xem xem nó có khả năng đá ứng được không, nếu không thể, nó sẽ chuyển các yêu cầu này đến các server thực.
- Theo www.stayinvisible.com: proxy server là một loại buffer giữa máy tính của bạn và các tài nguyên trên mạng internet mà bạn đang truy cập, dữ liệu bạn yêu cầu sẽ đến proxy trước, sau đó mới được chuyển đến máy của bạn.




**Hình 14 Mô hình proxy**

## 2.2 Tại sao proxy lại ra đời:

- Tăng tốc kết nối: các proxy có một cơ chế gọi là cache, cơ chế cache cho phép proxy lưu trữ lại những trang được truy cập nhiều nhất, điều này làm cho việc truy cập của bạn sẽ nhanh hơn, vì bạn được đáp ứng yêu cầu một cách nội bộ mà không phải lấy thông tin trực tiếp từ internet.
- Bảo mật: mọi truy cập đều phải thông qua proxy nên việc bảo mật được thực hiện triệt để.
- Filtering: ngăn cản các truy cập không được cho phép như các trang đồi trụy, các trang phản động…

## 2.3 Tổng kết chung về proxy:

- Theo các định nghĩa cũng như những giá trị mà proxy mạng lại như đề cập ở trên, ta có thể thấy proxy quả thật rất có lợi
- Tuy nhiên, lợi dụng về ý tưởng proxy, một số server trên mạng đã tự biến mình thành những trạm chung chuyển, những trung gian cho các kết nối không được cho phép. Chính điều này đã đưa ra thêm một định nghĩa mới, một ý nghĩa mới giành cho proxy.

- Rất nhiều địa chỉ trên mạng do một lý do nào đó mà bị cấm truy cập đối với người dùng như là các trang web đồi trụy, các trang phản động, nội dung không lành mạnh…. Tuy nhiên, để chống lại điều này, như đã nói ở trên, một số server đã biến mình thành proxy để giúp cho những kết nối cấm này có thể thực hiện được.
- Proxy này có 2 loại, hay nói cách khác là có 2 cách thông qua các proxy này để truy cập, đó là HTTP proxy và web-based proxy mà chúng ta sẽ được tìm hiểu ở phần sau. Và đây cũng chính là 2 phương pháp lập trình vượt firewall mà chúng em muốn nói đến trong luận văn này.

# Chương 3: CÁC PHƯƠNG PHÁP LẬP TRÌNH VƯỢT FIREWALL

## 3.1 Vượt firewall là gì:

- Nói một cách nôm na, vượt firewall là vượt qua sự truy cản của các chương trình bảo mật (Firewall) để có thể truy cập đến được đích mong muốn
- Vượt firewall có thể là vượt từ bên trong ra hay từ bên ngoài vào
- Ở đây, chúng ta chỉ đề cập đến vượt firewall từ bên trong ra, do đó chúng ta có thể tóm gọn lại có 3 hình thức vượt firewall: HTTP proxy, web-based proxy, http tunneling.

## 3.2 Phương pháp thứ nhất: HTTP Proxy

- Là phương pháp mà server sử dụng một cổng nào đó để trung chuyển các yêu cầu, các server này thường được gọi là web proxy server hay http proxy server
- Khi các yêu cầu của client bị từ chối bởi người quản trị (hay nói chính xác hơn là các chương trình quản lý trong mạng LAN), thì người sử dụng có thể sử dụng các proxy server để chuyển tiếp các yêu cầu mà trong đó, proxy server là một địa chỉ được cho phép kết nối đến.
- Các proxy server này thường không cố định, nó thường có thời gian sống rất ngắn.
- Sử dụng proxy này, bạn chỉ cần cấu hình mục proxy mà trong hầu hết các Web browser đều có hỗ trợ
- Phương pháp này sẽ được tìm hiểu sâu ở phần 2

## 3.3 Phương pháp thứ hai: Web-Based Proxy

- Phương pháp này cho phép người sử dụng truy cập vào các trang bị cấm dưới hình thức 1 truy cập vào 1 trang web trung gian.
- Đầu tiên người dùng truy cập vào trang web này
- Sau đó, người sử dụng cung cấp thông tin về trang web mà mình muốn đến (chủ yếu dưới hình thức url)
- Sau đó Web-base proxy này sẽ kết nối đến trang mà người dùng yêu cầu, lấy thông tin, đinh dạng lại thông tin, rồi gửi lại cho người dùng một cách hợp pháp
- Tất nhiên, web-based proxy này phải là một trang web mà chưa bị người quản trị cấm
- Phương pháp này sẽ được tìm hiểu sâu ở phần 2

## 3.4 Phương pháp thứ ba: Http Tunneling

- Cũng như các phương pháp trên, htttp tunneling cho phép người dùng truy cập vào những trang bị cấm
- Bao gồm một chương trình client ở phía người dùng và một chương trình ở phía server
- Đầu tiên, chương trình ở phía client sẽ tạo ra một đường hầm kết nối máy của bạn đến chương trình server đặt trên mạng, đường hầm này đi ngang qua firewall của bạn mà không hề hấn gì, vì địa chỉ server không bị filter. Khi đường hầm đã thiết lập xong mọi yêu cầu truy cập đến trang web sẽ thôn qua server, rồi đưa vào đường hầm và đến máy bạn mà firewall không hề hay biết. Do 1 số ứng dụng http-tunneling được viết theo mô hình client-server, cơ chế hoạt động dựa trên kịch bản làm việc dựng sẵn, ta có thể chủ động qua mặt các firewall bằng cách mã hóa các gói tin trao

- Do giới hạn của đề tài và giới hạn về mặt thời gian mà phương pháp này sẽ không được tìm hiểu kĩ trong luận văn.

**PHẦN THỨ HAI**

# VƯỢT FIREWALL

# Chương 4: VƯỢT FIREWALL BẰNG HTTP PROXY

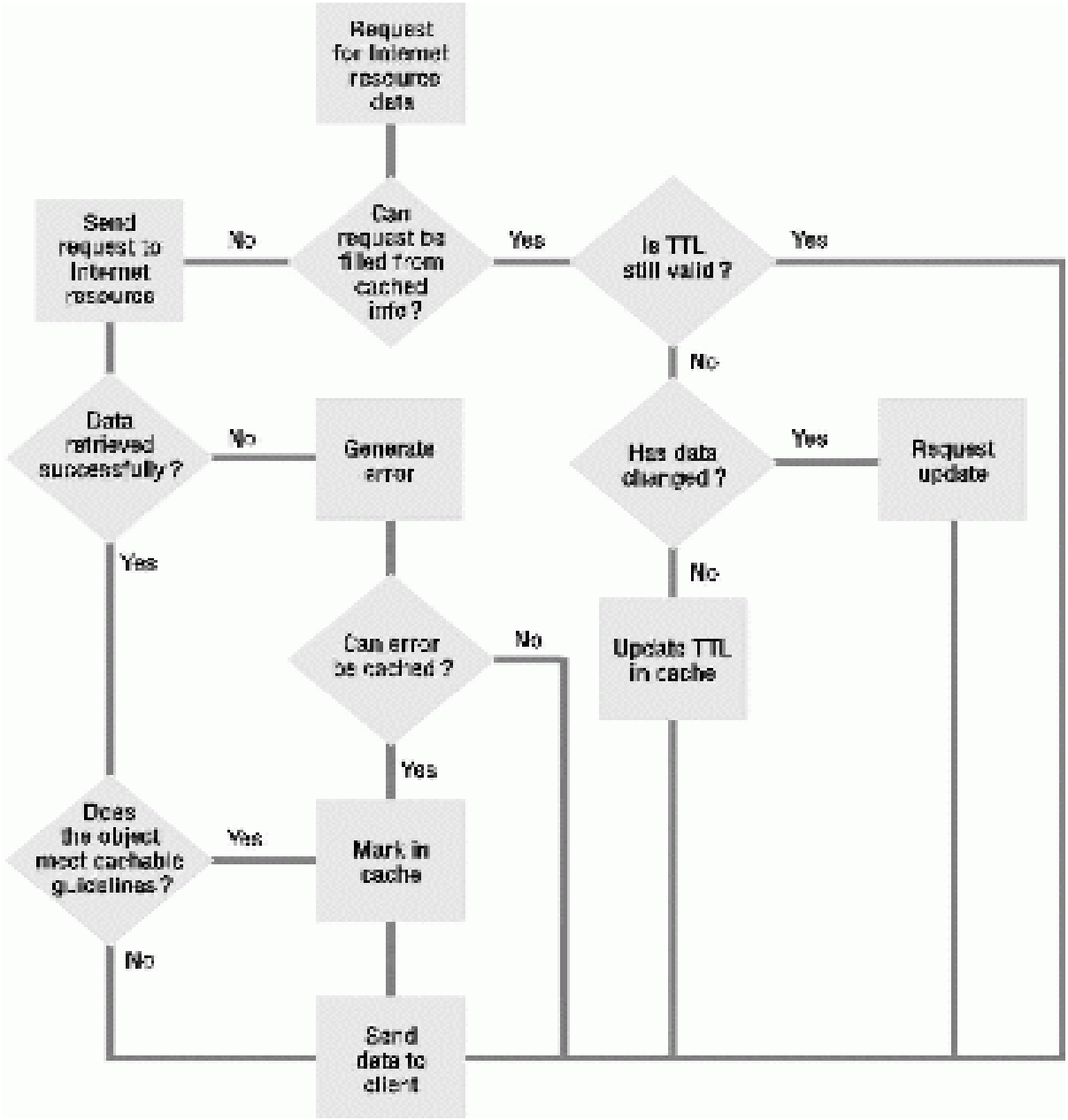
## 4.1 Khi các HTTP Proxy Server trở nên hữu ích:

- Nhiệm vụ chính của HTTP proxy server là cho phép những client bên trong truy cập ra internet mà không bị ngăn trở bởi Firewall (firewall). Lúc này tất cả các client phía sau Firewall đều có thể truy cập ra ngoài Internet chỉ với một chút công sức và không bị ngăn trở bởi các dịch vụ bảo mật
- Proxy server lắng nghe các yêu cầu từ các client và chuyển tiếp (forward) những yêu cầu này đến các server bên ngoài Internet. Proxy server đọc phản hồi (response) từ các server bên ngoài rồi gửi trả chúng cho các client bên trong.
- Thông thường, những client mà cùng subnet thì dùng cùng một proxy server. Do đó, proxy server có thể cache các document để phục vụ cho các client có cùng nhu cầu (cùng truy cập đến một trang chẳng hạn).
- Người dùng khi sử dụng proxy cảm thấy họ đang nhận các phản hồi một cách trực tiếp từ bên ngoài. Nhưng thực sự thì họ đang ra ngoài Internet một cách gián tiếp thông qua proxy.
- Các client mà không sử dụng DNS vẫn có thể duyệt web vì họ chỉ cần một thông tin duy nhất, đó là địa chỉ IP của proxy server. Tương tự, các cơ quan, doanh nghiệp… sử dụng các địa chỉ ảo (10.x.x.x, 192.168.x.x,

172.16.x.x → 172.32.x.x) vẫn có thể ra ngoài Internet một cách bình thường thông qua proxy server.

- Các proxy server có thể cho phép hay từ chối các yêu cầu dựa trên giao thức của các kết nối. Ví dụ như: một proxy server có thể cho phép các kết nối HTTP trong khi từ chối các kết nối FTP
- Khi bạn dùng proxy server như một cổng ra ngoài Internet từ mạng LAN, bạn có thể chọn lựa các tùy chọn như sau:

- Cho phép hay ngăn chặn client truy cập Internet dựa trên nền tảng địa chỉ IP
- Caching document: lưu giữ lại các trang web phục vụ cho các nhu cầu giống nhau
- Sàng lọc kết nối
- Cung cấp dịch vụ Internet cho các công ty dùng mạng riêng (nền tảng IP ảo)
- Chuyển đổi dữ liệu sang dạng HTML để có thể xem bằng trình duyệt

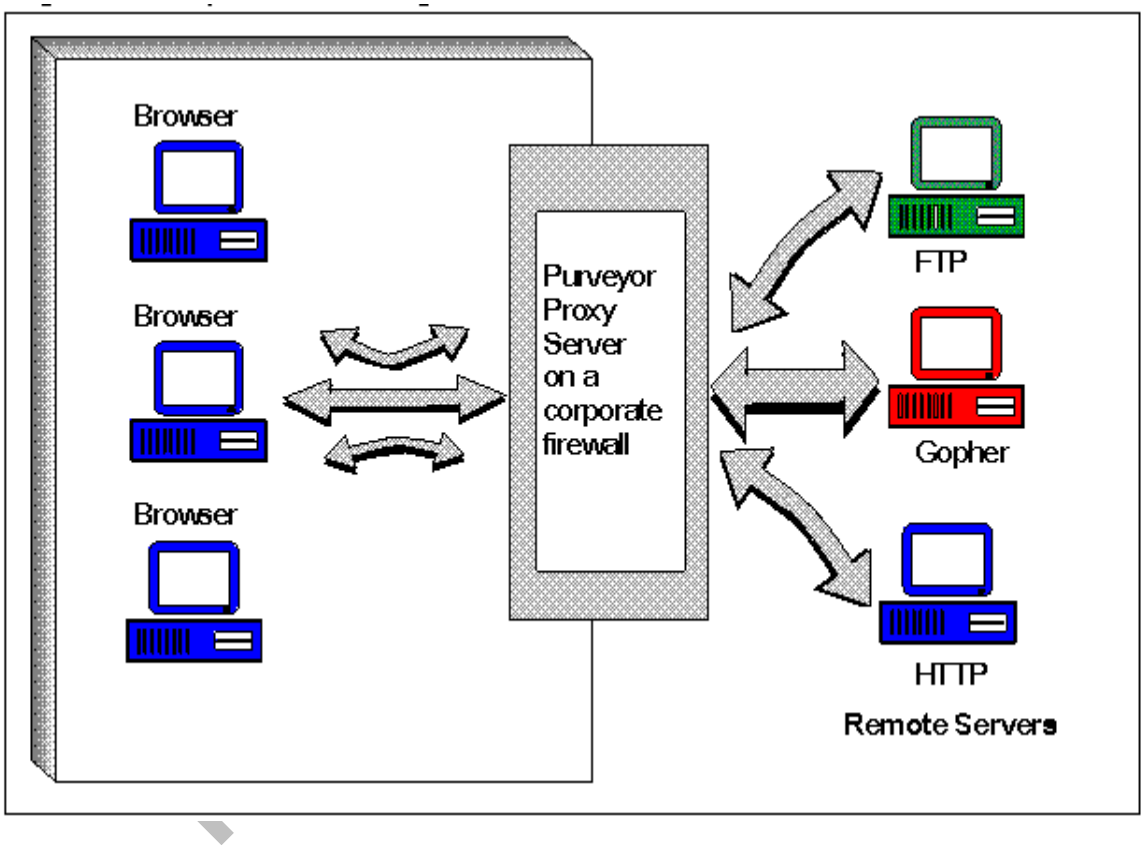



**Hình 15 Mô hình hoạt động chung của các proxy**

## 4.2 Chức năng chính:

### 4.2.1 ***Truy cập Internet:***

- Các máy trong mạng LAN có thể không thể truy cập đến các tài nguyên trên Internet một cách trực tiếp vì chúng đang hoạt động phía sau một bức Firewall. Trong trường hợp này, proxy server có thể giúp chúng thực hiện điều này một cách dễ dàng.



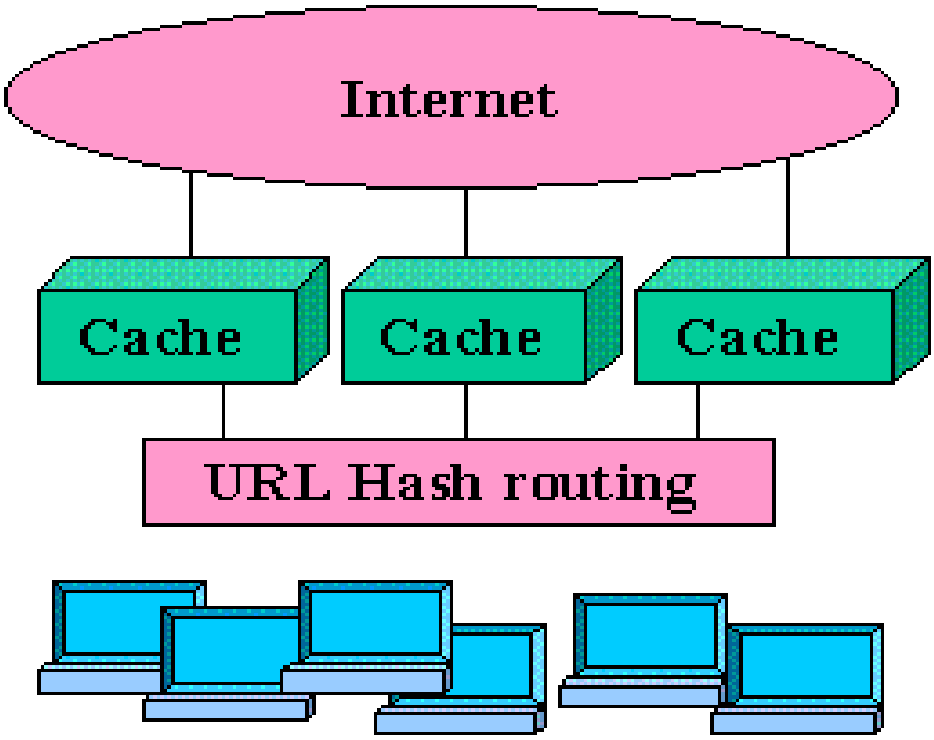

**Hình 16 Một số protocol được hỗ trợ**

- Ở hình trên, proxy server đang chạy trên một firewall host và thiếp lập các kết nối ra thế giới bên ngoài. Chúng ta cũng có thể sử dụng một máy tính khác để làm proxy server, máy này phải có đầy đủ các quyến truy cập Internet.

- Proxy nhận các yêu cầu từ trình duyệt, proxy truy vấn đến các thông tin được yêu cầu, chuyển đổi sang dạng HTML rồi gửi trả lại cho browser phía bên trong firewall. Proxy server có thể quản lý tất cả các kết nối ra ngoài Internet nếu nó là máy tính duy nhất có kết nối trực tiếp ra ngoài Internet.

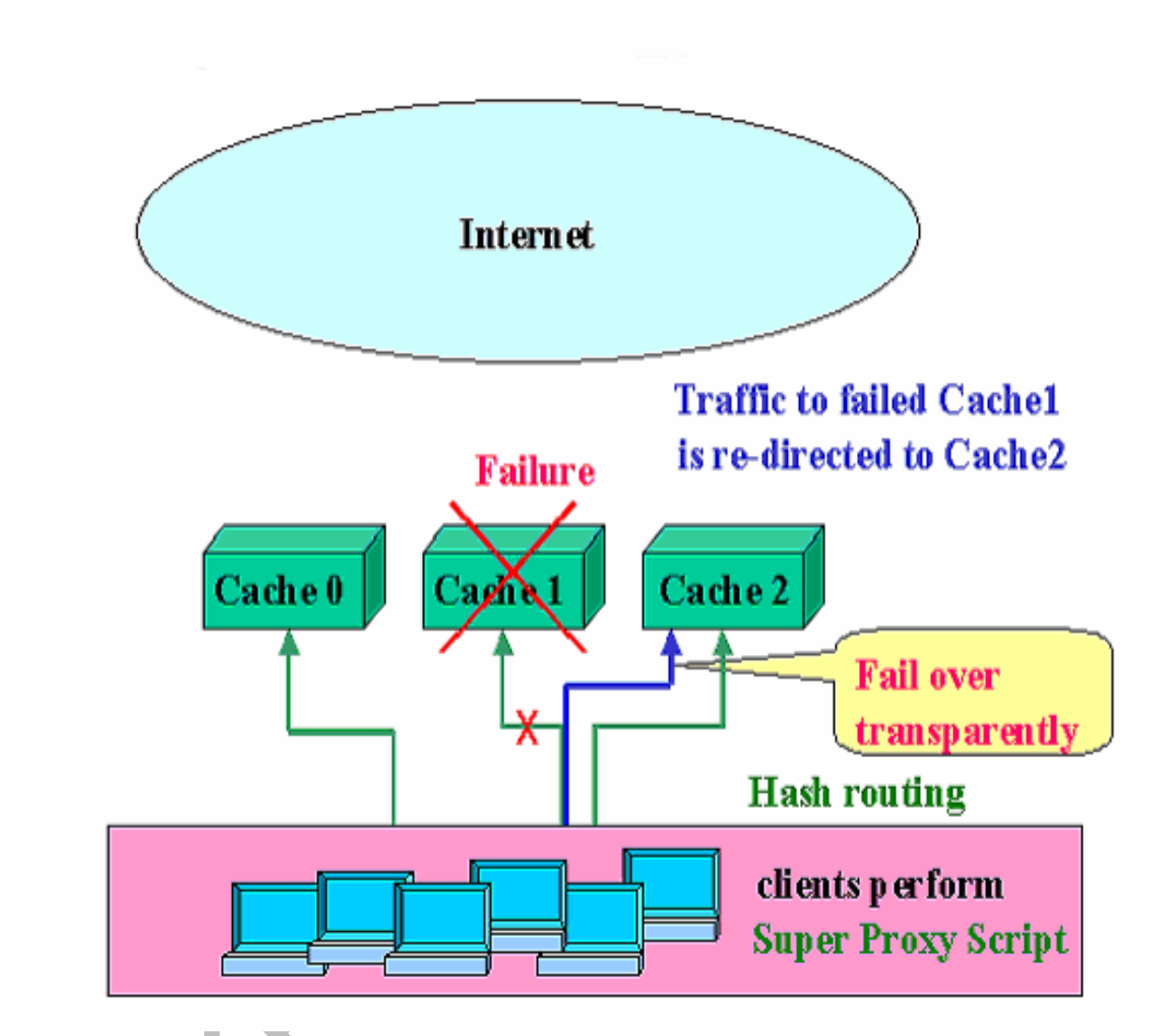
### 4.2.2 ***Caching documents:***

- Thông thường, các client của cùng một subnet truy cập đến một Web proxy server. Một vài proxy server cho phép bạn cache (lưu trữ tạm thời) các tài liệu này trên máy để phục vụ cho các máy khác có cùng nhu cầu. Giả sử: máy A vừa truy cập vào trang http://mail.yahoo.com , sau đó máy B lại yêu cầu đến trang này, trong trường hợp này, proxy server sử dụng lại documents này có sẵn trong máy mà không phải lên tận server lấy về. Điều này khiến cho tốc độ cải thiện rõ rệt




**Hình 17 Caching**

- Caching trên proxy server hiệu quả hơn trên máy đơn, nó sẽ tiết kiệm được không gian lưu trữ bởi vì bạn chỉ phải lưu lại một lần. Caching trên proxy server để cho hiệu quả hơn, chúng ta nên caching lại những trang mà thường xuyên được tham chiếu đến (được truy cập đến)
- Thông qua caching, chúng ta còn có thể truy cập đến trang đó ngay cả trong trường hợp server đó bị down
- Một số loại proxy cho phép cache ở nhiều nơi để đề phòng khi cache bị down hay bị lỗi




**Hình 18 Caching bị lỗi (failure)**

### 4.2.3 *Điều khiển truy cập Internet một cách có chọn lọc:*

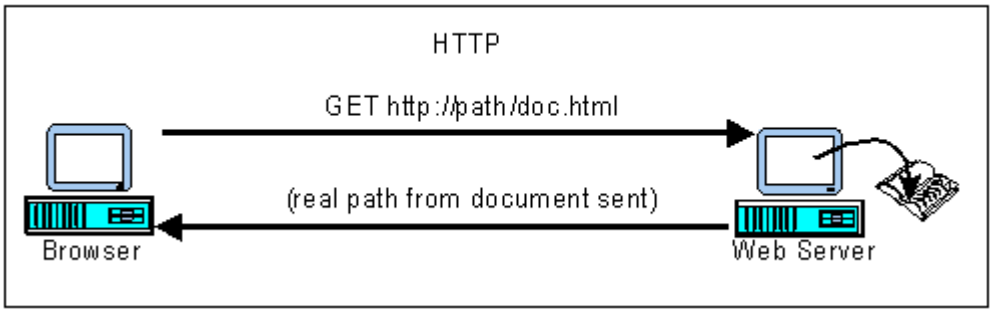
- Khi sử dụng proxy server bạn có thể lọc các transaction của các client. Một vài proxy server cho phép bạn:
  - Yêu cầu nào được chấp nhận, yêu cầu nào không
  - Ngăn chặn các trang mà bạn không muốn cho user truy cập đến

- Giới hạn các dịch vụ mà bạn muốn, ví dụ: bạn có thể cho phép user sử dụng dịch vụ HTTP nhưng lại không muốn cho họ sử dụng dịch vụ FTP

### 4.2.4 *Cung cấp dịch vụ Internet cho các cơ quan sử dụng IP ảo:*

Các tổ chức mà sử dụng một hay nhiều không gian địa chỉ ảo có thể sử dụng Internet, điều này hoàn toàn có thể. Bằng cách thông qua proxy server và proxy server sẽ giữ địa chỉ thật.

## 4.3 Một phiên giao dịch (transaction) thông qua proxy :




**Hình 19 Một transaction qua proxy**

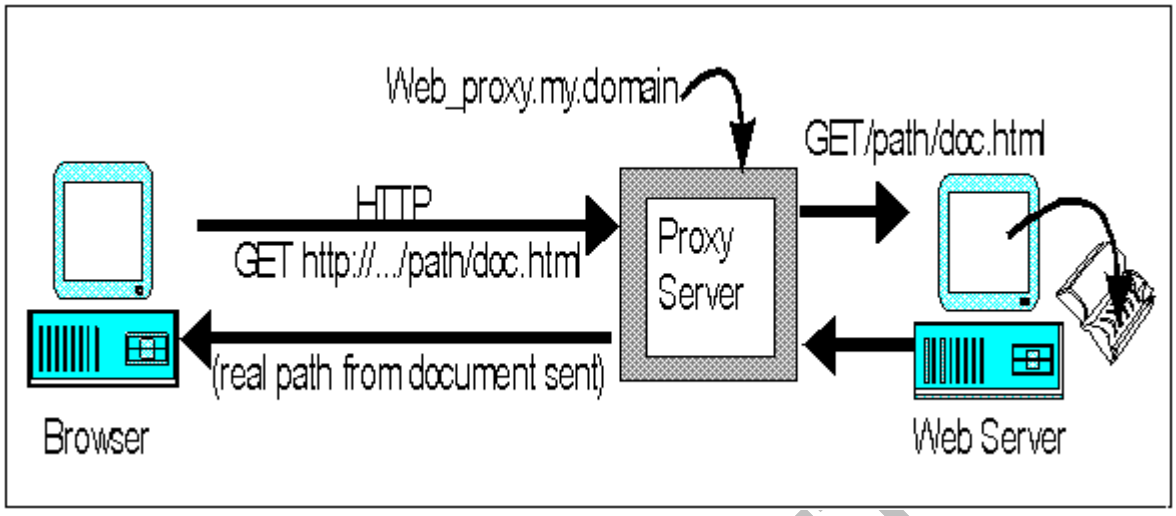
- Các client đều có các địa chỉ IP của nó cũng như một kết nối trực tiếp đến các server trên Internet. Khi trình duyệt tạo ra một yêu cầu HTTP thì HTTP server chỉ lấy đường dẫn và phần từ khóa của URL được yêu cầu, những phần khác như phần giao thức, hostname của máy đang chạy HTTP server đều đã rõ ràng đối với server.
- Ví dụ: khi bạn gõ: http://abc.com/class/th01.htm thì trình duyệt sẽ chuyển sang là: GET /class/th01.htm. Trình duyệt kết nối đến abc.com server, đưa ra lệnh và đợi phản hồi. Trong ví dụ này, trình duyệt tạo ra một yêu cầu đến HTTP server và chỉ rõ tài nguyên resource nào cần được tải về, không có giao thức cũng như không có bất kì hostname nào trong URL

## 4.4 Kết nối thông qua proxy server:

- Proxy server hoạt động với cả 2 vai trò là client và server, nó đóng vai trò server trong trường hợp nó tiếp nhận các yêu cầu HTTP từ các trình duyệt và hoạt động như một client khi nó kết nối đến server ở xa để truy vấn các tài nguyên
- Proxy sử dụng lại tất cả các thông tin mà trình duyệt đã gửi cho nó để gửi yêu cầu đến server ở xa nên sẽ không sợ bị mất mát hay thiếu hụt thông tin
- Một proxy server hoàn chỉnh có thể hỗ trợ hết tất cả các giao thức như: HTTP, FTP, Gopher, WAIS. Một proxy cũng có thể chỉ hỗ trợ một giao thức như HTTP nhưng điều đó thật bất tiện khi bạn có nhu cầu kết nối đến FTP trong quá trình bạn duyệt Web

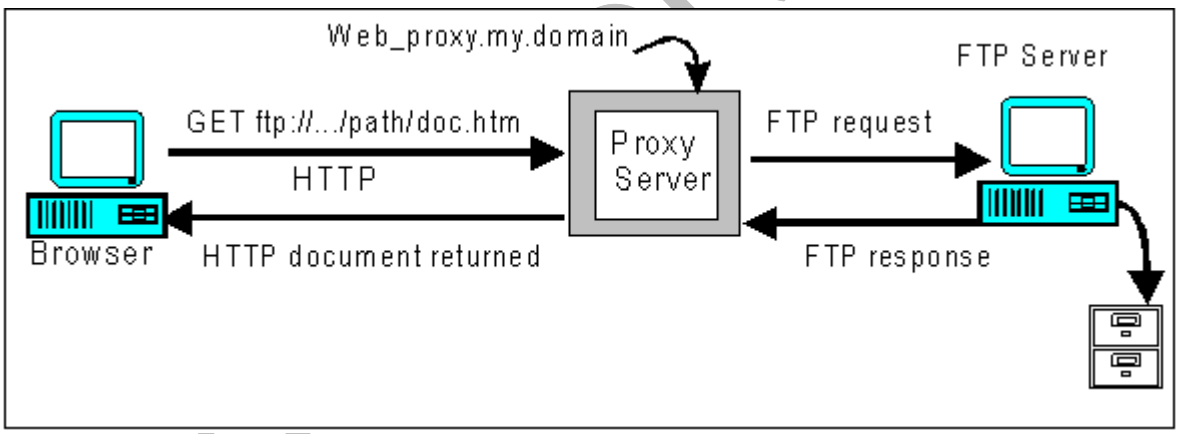
## 4.5 HTTP proxy:

- Khi proxy server đóng vai trò client, nó hoạt động như một trình duyệt nhận các resource.
- Một ví dụ về quá trình trao đổi thông tin:
  - Khi bạn gõ: http://abc.com/class/th01.htm
  - Trình duyệt chuyển URL này thành: GET http://abc.com/class/th01.htm
  - Yêu cầu này được đưa đến cho proxy server. Proxy server sẽ dựa vào URL tách lấy phần abc.com để kết nối đến remote server, sau đó chuyển URL thành: GET /class/th01.com , chuyển lệnh đến server rồi đợi phản hồi như hình bên dưới.




**Hình 20 Truy xuất thông tin thông qua HTTP proxy**

## 4.6 FTP proxy:




**Hình 21 Truy xuất thông tin thông qua FTP proxy**

Hình trên cho thấy quá trình một yêu cầu FTP thông qua proxy. Proxy server thông qua URL biết được đây là một yêu cầu FTP, do đó nó sẽ thực hiện một kết nối FTP đến server ở xa. Proxy server tạo một kết nối và truy vấn file đến FTP ở xa, lấy file về rồi gửi trả lại cho client.

## 4.7 Tiện lợi và bất tiện khi cache các trang Web:

- Caching có nghĩa là lưu trữ tài liệu trên máy cục bộ, vì vậy mà các user không phải kết nối đến server đế lấy các file về. Khi một trình duyệt cục bộ yêu cầu một file nào đó, proxy xem xét xem có có cache file đó lại không. Nếu có, nó sẽ gửi file về cho trình duyệt. Nếu bạn sử dụng tính năng này, bạn cần phải quyết định về:
  - Các trang nào cần được cache lại (tần số được truy cập nhiều)
  - Thời gian bao lâu phải cập nhật lại các trang này.
- Những thuận lợi của tính năng caching:
  - Caching tiết kiệm được một lượng lớn thời gian cho các user khi thường xuyên truy cập đến một trang nào đó. Proxy server sẽ đáp ứng các yêu cầu này một cách nhanh chóng vì chỉ phải truy vấn đến các file được lưu trữ cục bộ
  - Tiết kiệm được không gian lưu thông mạng
  - Tiết kiệm được không gian đĩa dùng để lưu trữ vì tất cả các máy cục bộ đều dùng chung một file thay vì các máy phải cache lại trên máy mình
  - Vẫn có thể cung cấp nhu cầu Internet ở một mức nào đó ngay cả khi không có kết nối Internet

## 4.8 Những bất cập do proxy:

- Tuy proxy như nói ở trên đem lại rất nhiều điều hữu ích. Tuy nhiên các gì cũng có 2 mặt và proxy cũng không ngoại lệ. Lợi dụng ý tưởng về proxy, hàng loạt các máy tính trên mạng tự biến mình thành những proxy server để cho các client có thể truy cập vào những trang có nội dung xấu mà nhà cung cấp dịch vụ đã ngăn chặn bằng firewall.
- Vấn đề được đặt ra là làm thế nào để cho các client truy cập Internet vẫn có thể truy cập Internet bình thường nhưng không thể truy cập những trang bị

chặn, hay nói cách khác là cấm cản người dùng sử dụng proxy bên ngoài hệ thống.

## 4.9 Kĩ thuật lập trình một HTTP Proxy cơ bản:

Lập trình một HTTP proxy cần qua các bước sau:

- Lắng nghe các kết nối đến proxy server
- Khi có kết nối đến thì tạo ra một thread để quản lý kết nối này
- Tiếp nhận và sửa đổi lại gói tin HTTP Request cho hợp lệ.
- Phân tích URL, lấy được phần tên trang Web và Port.
  VD:www.yahoo.com:8080 có tên là www.yahoo.com và port là 8080 (nếu không có giá trị port thì mặc định port=8080).
- Sử dụng phần tên này để phân giải địa chỉ lấy số IP.
- Kết nối đến remote server
- Chuyển yêu cầu đến server
- Chờ đợi thông tin phản hồi từ remote server
- Chuyển phần gói tin này về lại cho user.

# Chương 5: Vượt firewall bằng Web-Based Proxy

## 5.1 Thế nào là 1 web-based anonymous proxy ?

Web-based Anonymous Proxy là 1 dạng khác của Web Proxy Server, nhưng được xây dựng dưới dạng 1 trang web (tạm gọi là Web-based Proxy WBP) . Sau đây là các đặc điểm khác biệt của nó so với Web Proxy :

- Dễ dàng, thân thiện với người dùng do được Proxy tích hợp sẵn bên trong trang Web, người dùng chỉ cần cung cấp địa chỉ trang web cần đến (URL) cho WBP và bắt đầu duyệt web. Ngoài ra người dùng không cần phải tinh chỉnh các thông số khác địa chỉ IP của WBP, số hiệu cổng,.. cho trình duyệt của mình, chỉ cần biết tên hoặc IP của WBP và link đến WBP này
- Khi được các client yêu cầu, WBP sẽ lấy các thông tin (Resource) từ web server đích, sau đó xây dựng lại thành 1 trang web hoàn chỉnh rồi đẩy toàn bộ nội dung trang web hoàn chỉnh này về cho trình duyệt của Client. Thường thì trình duyệt phía Client sẽ nhận được trang web mình yêu cầu có đính kèm theo phần tiêu đề của WBP.
- Có khả năng chọn lọc các web page components khi được yêu cầu. VD: quyết định xem có cho phép sử dụng cookies,hình ảnh,javascript,cửa sổ pop-up,... trong trang web hay không.
- Do bản chất là "lướt web ẩn danh" thông qua 1 trang web trung gian nên các gói tin request của Client gần như giống hoàn toàn với các gói tin HTTP request thông thường .Vì vậy các phần mềm lọc gói tin sẽ khó lòng phát hiện ra đâu là gói tin "có vấn đề".
- Địa chỉ 1 số các WBP tham khảo khác trên internet :

- http://www.anonymization.net
- http://www.anonymizer.com

- http://www.stayinvisible.com
- http://www.proxify.com
- http://www.silentsuft.com

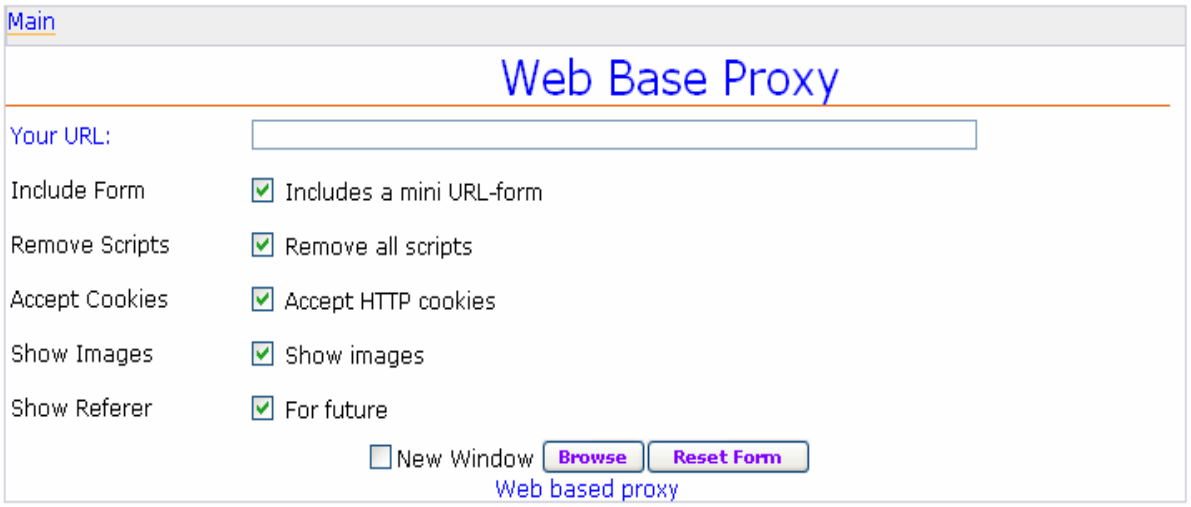
## 5.2 Cách thức hoạt động của 1 WBP :

Mỗi khi nhận được yêu cầu request từ phía Client,WBP sẽ :

- Phân tích URL để tiến hành tiếp nhận các resource tương ứng (links,hình ảnh,flash,…) từ trang web được client yêu cầu
- Sau khi nhận xong,WBP sẽ cập nhật lại các URLs của trang HTML được yêu cầu sao cho phù hợp. WBP sẽ tiến hành sàng lọc các thành phần (web page components) dựa theo yêu cầu Client và đẩy toàn bộ trang HTML đã được xây dựng lại này về phía Client
- Phía trình duyệt Client đang lắng nghe phản hồi từ phía WBP nên khi nhận được phản hồi, trình duyệt sẽ thể hiện trang web cho người dùng.

## 5.3 Giới thiệu về trang Web Based Proxy:

### 5.3.1 ***Giao diện:***

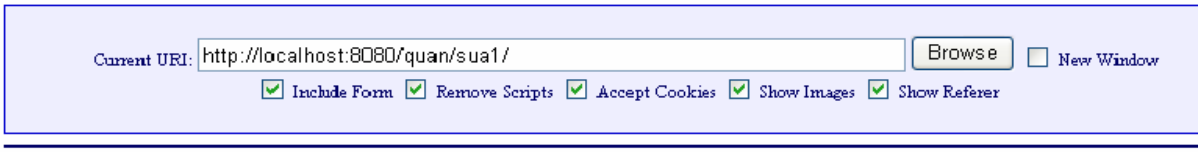



**Hình 22 Giao diện chính của Web Base Proxy**

- Trang web có giao diện đơn giản. Phía trên có một thanh textbox, cho phép user nhập địa chỉ trang web muốn đến
- Phía dưới là các option cho phép user lựa chọn
- Cuối cùng là 2 nút, cho phép người dùng kích hoạt cho trang web chạy và nút reset lại default.

### 5.3.2 ***Chức năng:***

- Cho phép người dùng nhập vào một địa chỉ dạng url. Người dùng chỉ cần nhập địa chỉ, bấm Enter, trang web sẽ tải nội dung mà người dùng muốn.
- Cho phép sử dụng các option, trong đó
  - Include a mini URL – form: thêm một phần của Web base Proxy vào đầu trang

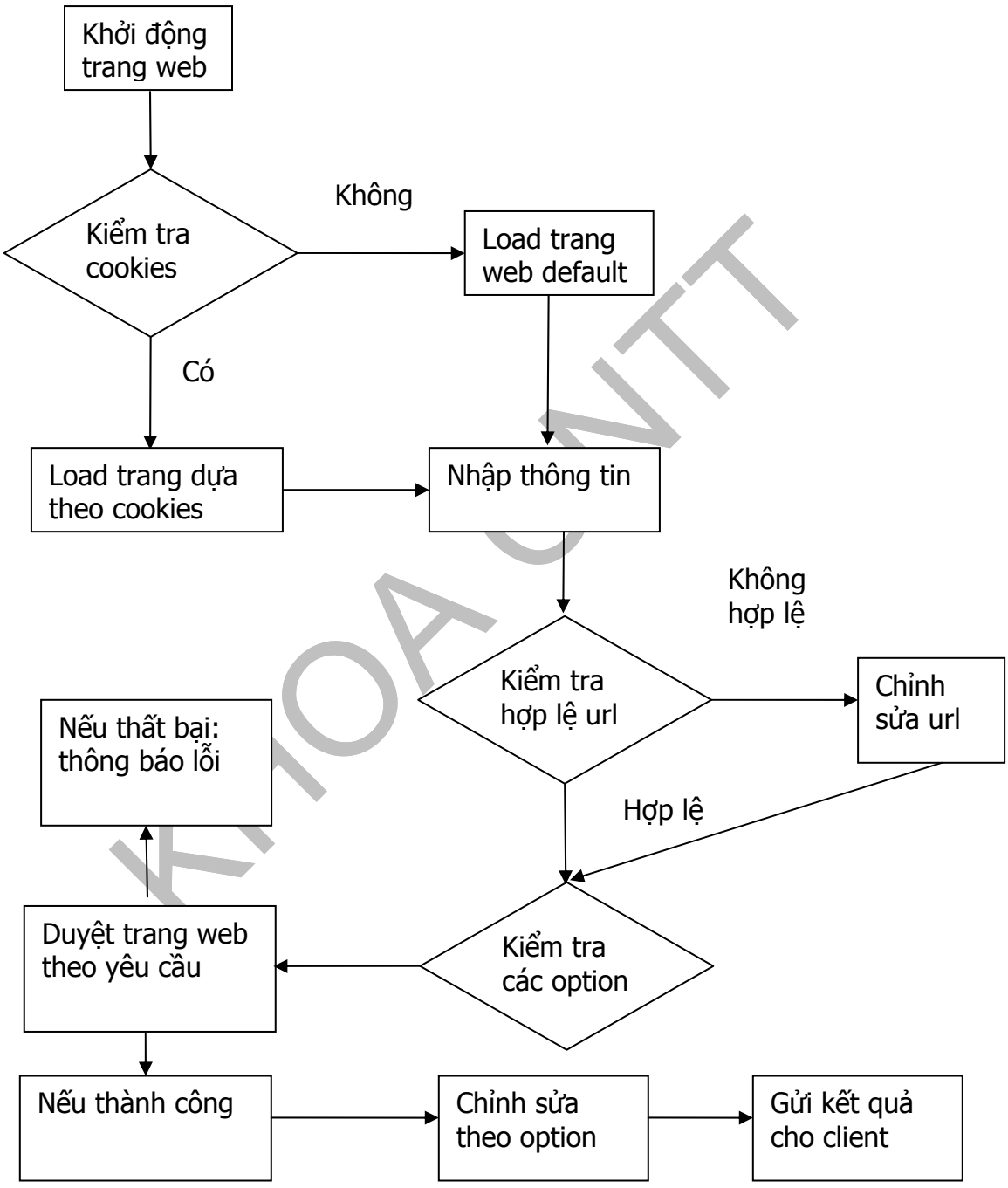



**Hình 23 Mini form trên mỗi đầu trang**

- Remove all scripts: Loại bỏ tất cả các script
- Accept HTTP cookies: cho phép sử dụng cookies để cải thiện tốc độ
- Show images: Tải nội dung trang web về trong đó có cả hình (lấy luôn hình, không loại bỏ)
- For future: Để dành cho tương lai
- New window: cho phép browse trong một cửa sổ mới.

### 5.3.3 ***Thuật toán:***

#### 5.3.3.1 Giới thiệu mô hình hoạt động:




**Hình 24 Sơ đồ hoạt động của 1 trang Web-Based Proxy**

#### 5.3.3.2 Diễn giải mô hình:

- Khởi động trang web: Bao gồm việc load các form, các đề mục, giao diện trang web
- Kiểm tra cookies:Kiểm tra xem trên máy hiện có sử dụng cookies của trang hay không
- Load trang web default:Nếu kiểm tra cookies không có, trình duyệt sẽ load trang mặc định, tức là url sẽ trống, các option mặc định sẽ được check…
- Load trang dựa theo cookies:Nếu kiểm tra cookies có, thì sẽ load theo cookies, bao gồm các url đã được sử dụng, các trạng thái của các option.
- Nhập thông tin:Client nhập các thông tin như url của trang web cần đến, check hay bỏ check các option tùy theo người dùng.
- Kiểm tra hợp lệ url:Kiểm tra về hình thức nhập như có thiếu http hay không, có thiếu www hay không, nếu thiếu sẽ tự động add thêm vào cho hợp lệ.
- Kiểm tra các option:Kiểm tra các option xem option nào được check, option nào không được check để thực hiện đúng theo yêu cầu của client.
- Duyệt trang web theo yêu cầu:Gửi yêu cầu đến webserver tương ứng: phân giải tên miền, gửi yêu cầu http đến server
- Thất bại, thông báo lỗi:Nếu không có trang web, địa chỉ sai do người dùng đánh sai hay bất cứ nguyên nhân nào làm cho việc gửi http request không được đáp ứng thì đều thông báo lỗi
- Thành công, chỉnh sửa theo option:Nếu thành công thì sẽ chỉnh sửa lại trang: dựa theo các option, xem có phải add thêm phần phụ

vào đầu trang hay không, lấy hay loại bỏ hình ảnh, lấy hay loại bỏ các script…(các mục này được thực hiện khi gửi http request).

- Gửi kết quả cho client:Gửi kết quả cuối cùng đến cho client là một trang web đã được tinh chỉnh lại, được chỉnh sửa lại cho phù hợp.

#### 5.3.3.3 Diễn giải một số hàm quan trọng :

- Hàm submit_form():Gửi yêu cầu đến server
- File url_form.inc:Phần header của trang gửi cho client.
- File style:Chứa các thông tin về giao diện: màu sắc, kích thước…
- Hàm set_response(): cấu trúc hóa lại trang web
- Hàm set_url(): Kiểm tra và tinh chỉnh url lại cho hợp lệ
- Hàm open_socket():Mở sock
- Hàm encode_url(): Mã hóa url
- Hàm decode_url(): Giả mã url
- Hàm set_flags(): Set các option
- Hàm set_cookies(): Ghi vào cookies
- Hàm get_cookies(): Lấy các thông tin từ cookies
- Hàm delete_cookies(): Xóa cookies
- Hàm include_form(): thêm form của web-base proxy vào phần đầu của trang (tùy thuộc vào option có được check)
- Hàm remove_scripts(): loại bỏ các script (tùy thuộc vào option có được check)
- Hàm send_response_headers(): gửi phần header cho client
- Hàm return_response():Gửi các phần còn lại cho client.
- Hàm remove_images():Loại bỏ các hình ảnh ra khỏi trang (tùy thuộc vào option có được check)

## PHẦN THỨ BA

# MODULE CHỐNG VƯỢT FIREWALL

Nội dung :

Do mục đích của luận văn là nghiên cứu các phương pháp lập trình vượt firewall nhằm tìm hiểu các cách thức mà người dùng có thể sử dụng để vượt qua firewal. Từ đó mở rộng ra xây dựng các module chống vượt firewall. Sau thời gian tìm hiểu, chúng em đã xây dựng được 2 module ứng dụng trên Windows nhằm ngăn chặn người dùng vượt firewall bằng 2 phương pháp đã trình bày bên trên :

- Module ứng dụng tích hợp vào trình duyệt Internet Explorer, nhằm phát hiện và ngăn chặn người dùng vượt firewall thông qua **Web based Proxy**. Module hoạt động dựa trên việc phân tích cách thức hoạt động của các trang Web-Based Proxy và đưa ra **3 chính sách để hình thành bộ lọc** cho Module. Khi người dùng duyệt bất kỳ 1 trang web nào, bộ lọc của module sẽ tiến hành kiểm tra dựa trên các chính sách đã được quy định sẵn, nếu vi phạm bất kỳ chính sách nào, trang web đó sẽ bị chặn lại và lưu thông tin (địa chỉ) vào cơ sở dữ liệu của module.
- Module ứng dụng dưới dạng 1 service trong hệ thống, nhằm phát hiện và ngăn chặn người dùng vượt firewall thông qua 1 **HTTP Proxy server**. Module bao gồm 2 phần chính: Lọc gói tin và chặn gói tin. Module hoạt động dựa trên việc lọc và kiểm tra nội dung các gói tin HTTP. Theo tài liệu RFC về HTTP, các gói tin **HTTP request thông qua 1 HTTP Proxy Server** sẽ có nội dung khác với các gói tin HTTP Request thông thường. Dựa trên đặc điểm này, module sẽ xây dựng chính sách lọc và kiểm tra các gói tin gửi đi trên Mạng. Khi 1 gói tin nào đó vi phạm, địa chỉ đích của gói tin đó (trường hợp này chính là địa chỉ IP của HTTP Proxy Server) sẽ được đưa vào bộ lọc và lưu vào cơ sở dữ liệu.

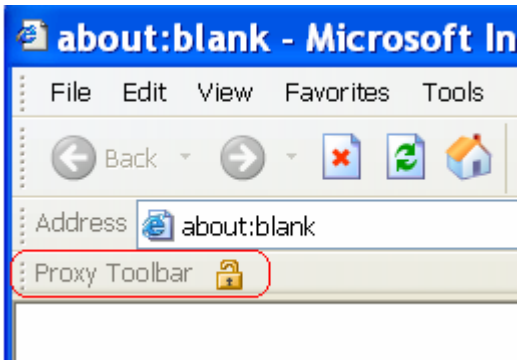
# Chương 6: Plug-in chống vượt firewall cho trình duyệt Internet Explorer

Chương này chúng em xin phép được trình bày về module thứ nhất: Plug-in chống vượt firewall cho trình duyệt Internet Explorer

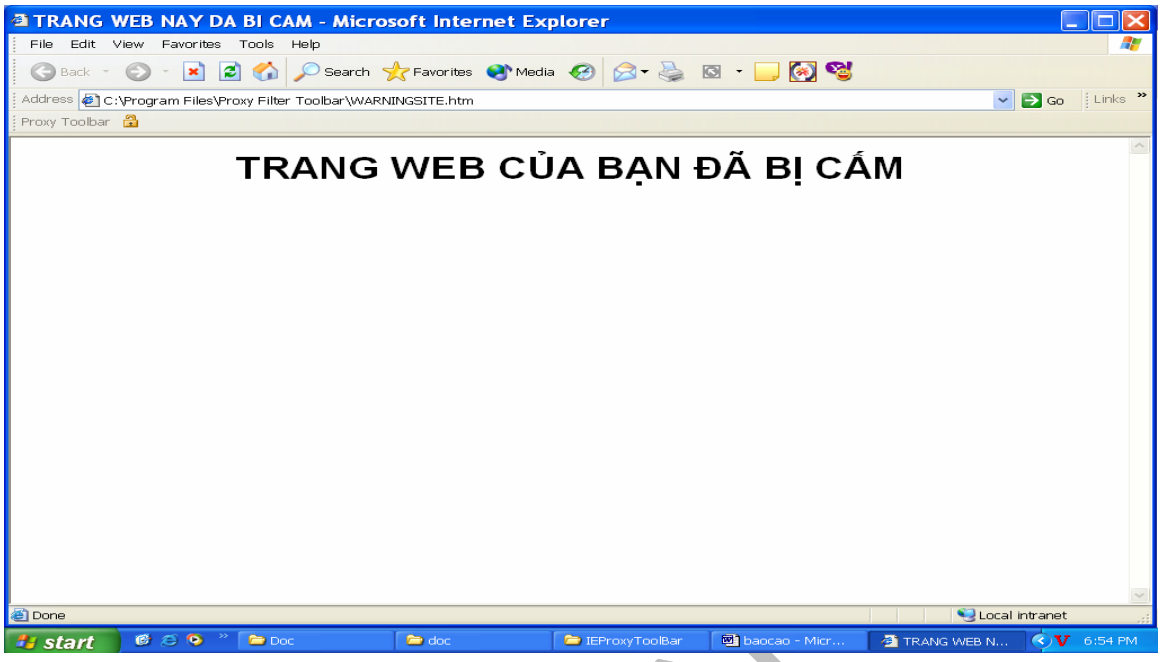
## 6.1 Giới thiệu sơ lược :

Plugin là 1 ứng dụng được viết tích hợp trong trình duyệt web Internet Explorer, có nhiệm vụ kiểm soát người dùng khi duyệt web. Nếu phát hiện người dùng có ý định muốn vượt qua firewall thông qua 1 trang Web Based Proxy nào đó, plugin sẽ tiến hành ngăn chặn và lưu thông tin về trang web này (địa chỉ trang web) vào cơ sở dữ liệu để làm cơ sở lọc về sau. Ứng dụng được viết trên môi trường Visual C 6.0 dưới dạng ATL, chạy tốt trên các phiên bản trình duyệt IE5 trở lên và các phiên bản từ Windows 2000 trở lên. Do nhu cầu lưu trữ thông tin về danh sách các Proxy Server, Web-based proxy làm cơ sở cho bộ lọc nên các thông tin này được module lưu trữ vào cơ sở dữ liệu Microsoft Access.

Giao diện chính của plugin




**Hình 25 Giao diện chính của plug-in**




**Hình 26 Trang thông báo mỗi khi người dùng duyệt những trang web vi phạm**

## 6.2 Các tính năng chính:

### 6.2.1 *Lọc các trang web dựa trên việc duyệt danh sách các trang web có sẵn trong cơ sở dữ liệu:*

Nếu người dùng có ý định muốn duyệt 1 trang web có địa chỉ đã được lưu trong cơ sở dữ liệu, plugin sẽ hiện ra trang thông báo người dùng đã bị cấm.

### 6.2.2 *Lọc các trang web dựa trên cơ chế kiểm tra địa chỉ (URL):*

Khi người dùng duyệt đến 1 trang web mới, nếu trang web này có thể giúp người dùng qua mặt được firewall (hay còn gọi là "vi phạm"), plugin sẽ hiện ra trang thông báo cho người dùng và lưu lại địa chỉ trang web này vào cơ sở dữ liệu. Do đại đa số các trang Web-based Proxy khi hoạt động thì thể hiện địa chỉ của mình dưới dạng *http://domain_name của WebProxy/địa chỉ thật của trang web muốn duyệt* nên dựa vào cơ chế này, ta có thể xác định các "địa chỉ

*Ví dụ:* giả sử trang web www.abc.com là 1 trang vi phạm

Khi người dùng thông qua trang này lướt vào nhưng trang mình muốn đến www.yahoo.com thì kết quả URL của trang này sẽ thê hiện như sau:

http://www.webproxy.com/www.yahoo.com

hay http://www.webproxy.com?url=www.yahoo.com

.....

Ta có thể dễ dàng tách địa chỉ trên ra làm 2 địa chỉ riêng biệt. Nếu gặp những địa chỉ quá rõ ràng như thế này thì bộ lọc chắc chắn sẽ phát hiện ra được và lưu địa chỉ mới này vào cơ sở dữ liệu cho những lần duyệt tiếp theo.

### 6.2.3 *Lọc dựa trên nội dung của các Input Form trong trang web:*

Trong trường hợp các trang tiến hành mã hóa địa chỉ hay thậm chí không thể hiện địa chỉ ra trình duyệt thì sao ???

Lúc này chức năng thứ 3 của bộ lọc lại trở nên hữu ích. Đây là 1 chức năng bổ sung cho trường hợp 2 nêu trên. Khi người dùng truy cập vào các trang Web-Proxy để truy cập vào các trang web khác thì gần như luôn luôn phải nhập địa chỉ trang web mình muốn đến vào 1 textbox, sau đó tiến hành submit cho webserver xử lí.

VD: 1 trang web-based proxy thường có cách trình bày như sau


**Hình 27 Cách trình bày thông thường của một trang web base proxy**

Có thể thấy được khi người dùng gõ đầy đủ tên trang web và click vào nút Go.Trang web sẽ submit nội dung text field vừa được nhập (http://www.google.com) lên cho server và server tiến hành duyệt

Dựa trên hành động này, bộ lọc sẽ tiến hành lọc các Input tag của trang web và kiểm tra xem có Input tag nào vi phạm hay không. Nếu vi phạm tức là gần như người dùng đang có ý định muốn submit 1 URL đến cho server và muốn truy cập đến trang này.

### 6.2.4 *Cập nhật các trang web based proxy:*

Cho phép người dùng có thẩm quyền được cập nhật (thêm xóa) danh sách các trang web based proxy trong cơ sở dữ liệu.

### 6.2.5 *Vô hiệu hóa/kích hoạt plugin:*

Cho phép người dùng có thẩm quyền được vô hiệu hóa/kích hoạt plugin.

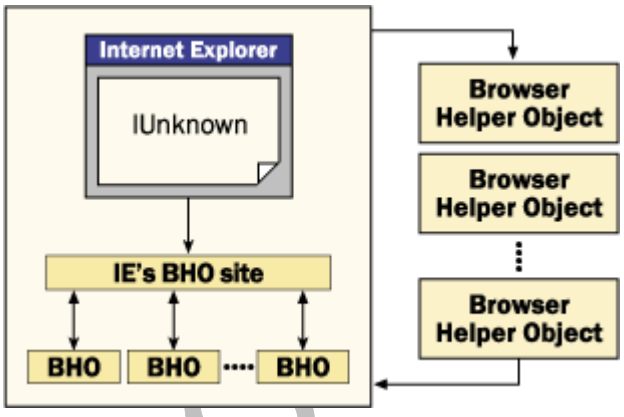
## 6.3 Một số vấn đề cần lưu ý khi viết plugin cho trình duyệt IE :

### 6.3.1 *Khái niệm Browser Helper Objects (BHO):*

**Browser Helper Objects (BHO)**, tạm dịch là đối tượng trợ giúp cho trình duyệt, là 1 khái niệm do **Microsoft** đưa ra. Đây là 1 dạng ứng dụng được phát triển dựa trên môi trường **COM** (Component Object Model). Dòng đời của đối tượng này gắn liền với dòng đời của trình duyệt Internet Explorer, tức là khi khởi động sẽ sử dụng chung vùng nhớ cùng với trình duyệt web Internet Explorer và chỉ được hủy khi trình duyệt bị đóng. Khi chạy, đối tượng sẽ có thể tương tác với tất cả mọi thành phần cũng như đối tượng khác của trình duyệt (ví dụ: cửa sổ, toolbar, textfield,…),có thể nhận được các thông điệp, sự kiện do trình duyệt phát ra như các sự kiện trở về trang trước đó (GoBack), trang sau

(GoForward), hay sự kiện Download thành công (DocumentComplete),… Các BHO khi được khởi tạo thì trước hết phải trải qua quá trình đăng kí vào Registry cho hệ thống thông qua giá trị của CLSID. Giá trị này đóng vai trò như 1 giá trị định danh (Identifier) cho duy nhất BHO.

Hình dưới đây minh họa quá trình trình duyệt khởi động và nạp các BHO vào bộ nhớ để xử lí:




**Hình 28 Quá trình trình duyệt khởi động và nạp các BHO**

Quá trình hoạt động như sau :

- Khởi động trình duyệt.
- Trình duyệt sẽ tìm trong Registry các giá trị CLSID của các BHO tương ứng và load các module ứng dụng của các BHO này vào bộ nhớ
- Mỗi BHO được khởi tạo sẽ có 1 Interface (tạm dịch là đối tượng giao tiếp) riêng biệt. Khi tìm thấy các Interface này của BHO, trình duyệt sẽ chuyển con trỏ trỏ đến Interface của chính mình (Interface **IUnkown**) cho các BHO. Chính việc chuyển **IUnkown** cho các BHO mà các BHO này mới có thể can thiệp được vào các đối tượng cũng như các sự kiện của trình duyệt.

### 6.3.2 ***Một số hàm xử lí quan trọng:***

- **HRESULT SetSite(IUnknown* pUnkSite)**
- Đây chính là hàm khởi tạo đối tượng BHO. Nhiệm vụ chính của hàm này là nhận con trỏ đối tượng IUnkown và 1 số đối tượng quan trọng khác (IWebBrowser2, IConnectionPointContainer) từ trình duyệt và lưu lại để xử lí.
- **HRESULT Connect(void)**
- Báo cho trình duyệt biết rằng BHO có ý định muốn bắt các sự kiện và xử lí trước khi gửi trả lại cho trình duyệt.
- **HRESULT Invoke()**
- Bắt các sự kiện do trình duyệt phát ra và chuyển đến hàm xử lí sự kiện tương ứng.
- **HRESULT Disconnect(void)**
- Khi đối tượng bị hủy hay chủ động kết thúc, cần gọi sự kiện này để thông báo chấm dứt việc xử lí các sự kiện cho trình duyệt
- **Các hàm xử lí sự kiện:** Tùy theo loại sự kiện mà BHO sẽ có các xử lí tương ứng, các sự kiện được xử lí trong Module này lần lượt là:
- **DISPID_BEFORENAVIGATE2**: Sự kiện chuẩn bị duyệt đến 1 trang web khác trang hiện hành.
- **DISPID_ONQUIT** : Sự kiện đóng trình duyệt
- Nói thêm về việc đăng kí BHO vào registry cho trình duyệt
- Mặc dù khi tạo 1 ứng dụng dạng COM Plugin cho Internet Explorer, Visual C++ 6.0 sẽ tự tạo các dòng lệnh khởi tạo các thông số cho ứng dụng trong regsitry trong **tập tin có đuôi là rgs**. Tuy nhiên các dòng lệnh đăng kí ứng dụng vào Registry thì người dùng phải tự thêm vào. Nội dung cần thêm vào như sau :

HKLM{SOFTWARE{Microsoft{ Windows {CurrentVersion {Explorer{'Browser Helper Objects'{ForceRemove {**Số ID đã được VC tạo sẵn**} = s '**Tên đối tượng BHO muốn thể hiện**'}}}}}}}

## 6.4 Chi tiết lưu trữ dữ liệu :

### 6.4.1 ***Bảng Forbidden***

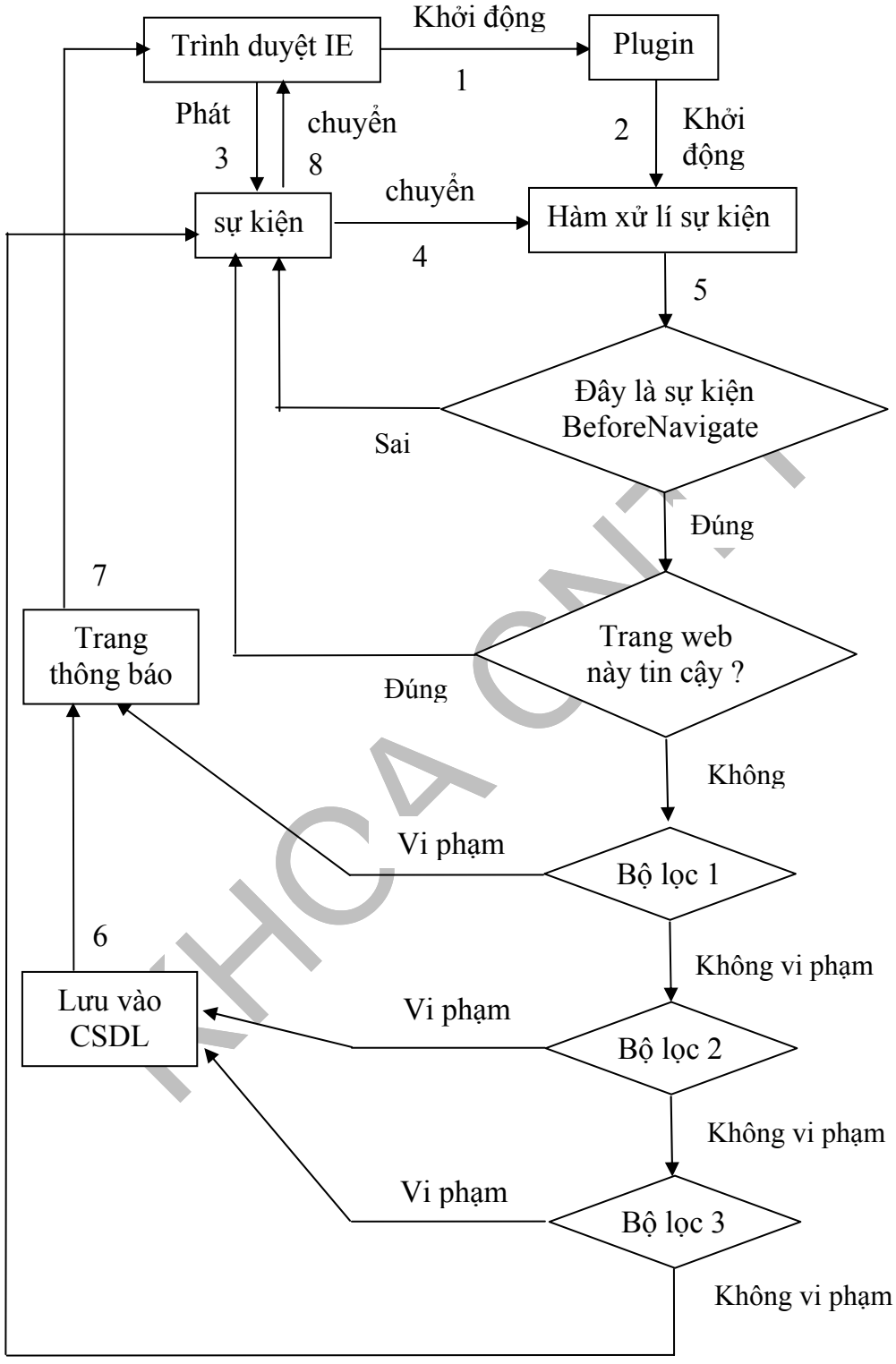
| Tên trường | Kiểu | Chú thích |
|---|---|---|
| URL | Text | Địa chỉ trang web based proxy bị cấm |

### 6.4.2 ***Bảng Trusted***

| Tên trường | Kiểu | Chú thích |
|---|---|---|
| URL | Text | Địa chỉ trang web tin cậy |

## 6.5 Thuật toán chính của ứng dụng :

### 6.5.1 ***Mô hình hoạt động của Plugin :***




**Hình 29 Mô hình hoạt động của Plugin**

### 6.5.2 ***Diễn giải mô hình :***

- **BeforeNavigate:** Sự kiện do trình duyệt phát ra khi người dùng chuẩn bị duyệt đến 1 trang web mới nào đó (khác với trang hiện hành). Ví dụ: Khi click chuột vào 1 link, 1 nút trên trang web và chuyển sang 1 trang web mới, khi gõ địa chỉ vào thanh address bar để chuẩn bị duyệt,…
- **Bộ lọc 1:** Nhận vào địa chỉ trang web không đáng tin cậy và tiến hành kiểm tra. Bộ lọc sẽ truy xuất vào cơ sở dữ liệu để duyệt xem trang web này có nằm sẵn trong danh sách các trang bị cấm hay không. Nếu có thì bộ lọc sẽ lưu địa chỉ này vào cơ sở dữ liệu và chuyển hướng đến trang thông báo cấm cho người dùng. Nếu không thì sẽ chuyển đến bộ lọc tiếp theo.
- **Bộ lọc 2:** Nhận vào địa chỉ trang web không đáng tin cậy và tiến hành kiểm tra. Nếu địa chỉ này chứa thêm 1 địa chỉ trang web khác thì được xem như vi phạm (đã trình bày ở trên). Bộ lọc sẽ lưu địa chỉ vi phạm này vào cơ sở dữ liệu.
- **Bộ lọc 3:** Nhận vào con trỏ đối tượng IWebBrowser2 để xử lí. Con trỏ này đại diện cho trang web hiện hành cần kiểm tra. Dựa vào con trỏ đối tượng này, ta có thể lấy được toàn bộ nội dung trang web (các thẻ HTML, các script,….). Như đã trình bày ở trên, bộ lọc 3 hoạt động dựa trên việc kiểm tra nội dung các INPUT FIELD của trang web. Do đó bộ lọc chỉ chú trọng đến việc lọc các thẻ INPUT của trang HTML. 1 trang web được bộ lọc xem là 1 trang Web Based Proxy khi và chỉ khi nó **chứa không quá 4 thẻ INPUT dạng text**, và **ít nhất 1 trong các thẻ Input này có nội dung là địa chỉ 1 trang web** nào đó. Nếu trang web nào thỏa điều kiện nêu trên thì sẽ được xem là vi phạm và lưu lại vào cơ sở dữ liệu.

## 6.6 Những ưu điểm và hạn chế:

Plugin áp dụng 1 số thuật giải Heuristic nhằm phát hiện các trang Web-proxy mới hoặc chưa có trong cơ sở dữ liệu, khiến bộ lọc thông minh hơn do có khả năng tự học các địa chỉ trang web mới muốn qua mặt Firewall. Khi chạy thử các trang web-proxy mới, bộ lọc hoạt động khá hiệu quả và chính xác.

Đa số các thuật giải này được xây dựng dựa trên việc quan sát quá trình vận hành của các trang Web-based Proxy và tìm ra những điểm chung đặc trưng và khác biệt so với các trang web khác làm cơ chế hoạt động cho bộ lọc. Do thuật giải không đảm bảo tính chính xác 100% nên có 1 số trường hợp thiếu sót hay thậm chí sai sót ngoài ý muốn. Đa số những sai sót đều rơi váo trường hợp khi người dùng sử dụng search engine (như google,yahoo,…) để tiến hành tìm kiếm 1 địa chỉ nào đó trên internet. Trong những trường hợp này, plugin sẽ tự cho rằng các trang web tìm kiếm này là các Web-Based Proxy và tiến hành ngăn chặn. Lỗi trên có thể khắc phục được bằng cách thêm vào danh sách các trang web tin cậy và buộc bộ lọc kiểm tra các "trang web tin cậy này" trước khi lọc. Tuy nhiên cách này cũng không thể khắc phục hoàn toàn.

Quá trình hoạt động của Plugin phụ thuộc khá nhiều vào "sự tồn tại" của tập tin cơ sở dữ liệu lưu trữ các trang Web-Based Proxy. Nên khi tập tin trên không tồn tại hay bị lỗi, tính năng lọc của Plugin chắc chắn không thể hoạt động chính xác được.

Trong quá trình chạy thử và kiểm lỗi, chúng em đã cố gắng sửa chữa hầu hết các sai sót này.Chúng em xin cố gắng phát triển thêm để bộ lọc ngày càng hoàn thiện hơn.

# Chương 7: SERVICE CHỐNG VƯỢT FIREWALL

Chương này chúng em xin phép được trình bày về module thứ hai: Service chống vượt firewall cho hệ điều hành Windows.

## 7.1 Giới thiệu sơ lược :

Service chống vượt Firewall là 1 ứng dụng được viết dựa trên mô hình Service truyền thống của Windows. Service là 1 ứng dụng chạy nền trong hệ thống, hoàn toàn tuân thủ theo các yêu cầu và tính năng bảo mật do Windows quy định (Chỉ có người chủ Service – trường hợp module này là admin hệ thống – mới có quyền tắt/mở/xóa serive mà thôi). Service chịu trách nhiệm lọc và bắt các gói tin gửi ra mạng ngoài (Internet) nhằm **phát hiện và ngăn chặn các gói tin gửi đến các HTTP Proxy Server** và lưu lại địa chỉ các HTTP Proxy Server này để làm cơ sở hoạt động cho bộ lọc. Service bao gồm 2 module nhỏ: **module bắt gói tin** và **module chặn địa chỉ IP**.

## 7.2 Các tính năng chính của module:

Theo như đã giới thiệu, module này được chia ra làm 2 module nhỏ riêng biệt, hỗ trợ nhau trong quá trình Service hoạt động: Đó là **module bắt gói tin** và **module chặn địa chỉ IP**.

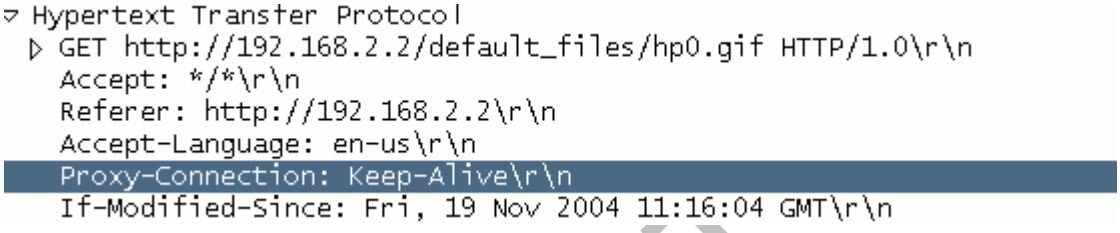
- *Module bắt gói tin:* module được viết dựa trên thư viện **Winsock2.0** của Windows, nhiệm vụ bắt các gói tin lưu thông ra/vào card mạng của hệ thống.
- *Module chặn địa chỉ IP:* module được viết dựa trên **mô hình Filter-Hook Driver** được Microsoft giới thiệu trong **tài liệu Windows 2000 DDK**. Ứng dụng viết dựa trên mô hình này có thể lọc các gói tin ra vào card mạng của hệ thống (theo tài Windows 2000 DDK). Theo tài liệu RFC về HTTP Protocol, các gói tin gửi đến HTTP Proxy Server đều có điểm đặc trưng riêng so với các gói tin khác. Service dựa vào đặc điểm này làm cơ sở hoạt động cho bộ lọc của mình.

## 7.3 Module bắt gói tin :

Module chịu trách nhiệm bắt và kiểm tra nội dung gói tin ra/vào card mạng.

### 7.3.1 ***Đặc điểm của gói tin HTTP request đến HTTP Proxy Server:***

Theo tài liệu RFC về HTTP Protocol, gói tin HTTP request đến Proxy server sẽ có định dạng như sau :




**Hình 30 Định dạng của gói tin gửi đến proxy server**

Trong hình minh họa trên, ta thấy nội dung 1 gói tin HTTP Request (câu lệnh HTTP ở đây chính là lệnh GET) được bổ sung thêm trường Proxy-**Connection: Keep-Alive**. Đây chính là đặc điểm mấu chốt để phận biệt gói tin HTTP Request đến 1 Proxy Server so với các gói tin thông thường khác.

### 7.3.2 ***Tóm tắt các bước cần lưu ý khi xây dựng module;***

- Khởi tạo các thông tin cần thiết (địa chỉ,port,..) cho 1 SOCK_RAW Socket.
- Chuyển chế độ hoạt động của Socket sang chế độ SIO_RCVALL (bắt tất cả các gói tin ra/vào hệ thống).
- Bắt đầu nhận và xử lí gói tin. Lưu ý: Do mục tiêu đề ra ban đầu của module là bắt và xử lí các gói tin HTTP (TCP) nên cần phải gỡ bỏ các Header của gói tin nhận được (đây là các gói IP) rồi mới bắt đầu xử lí.
- Tham khảo thêm tài liệu về cấu trúc gói tin TCP/IP và HTTP Protocol trong quá trình xử lí các gói TCP.

### 7.3.3 ***Chi tiết các đối tượng, hàm xử lí chính của module :***

- socket(AF_INET**, SOCK_RAW,** IPPROTO_IP)
  - Hàm tạo Socket. Lưu ý phải khởi tạo socket dạng SOCK_RAW thì mới có thể bắt được gói tin tầng IP.
- WSAIoctl(SOCKET s,DWORD *dwIoControlCode*, , , , , , ,)
  - Hàm thiết lập chế độ hoạt động cho socket. Chỉ cần lưu ý đến 2 tham số đầu tiên: SOCKET cần thiết lập và chế độ hoạt động. Ở đây *dwIoControlCode* phải bằng **SIO_RCVALL** thì module mới có thể bắt được các gói tin ra/vào card mạng
- Một số hàm liên quan khác: recv, WSAStartup, …

## 7.4 Module chặn địa chỉ IP:

Module chịu trách nhiệm lọc và chặn các gói tin ra/vào card mạng dựa trên địa chỉ IP. Module được xây dựng dựa trên mô hình Filter-Hook Driver của Windows 2000 DDK.

### 7.4.1 ***Giới thiệu về Filter-Hook Driver :***

Filter-Hook Driver là khái niệm được Microsoft đưa ra trong tài liệu về Windows 2000 DDK. Đây là Driver mở rộng các tính năng của IP Filter Driver (Có sẵn trong hệ điều hành Windows 2000 trở về sau).

Thực chất Filter-Hook Driver không phải là 1 trình điều khiển dành cho môi trường mạng, nó được xem như 1 trình điều khiển dành cho nhân của hệ thống (Kernel Mode Driver). Bên trong trình điều khiển này, chúng ta chỉ cần định nghĩa 1 hàm CALLBACK (1 dạng hàm bắt sự kiện) và đăng kí hàm CALLBACK này cho trình điều khiển bộ lọc địa chỉ IP của hệ thống (IP Filter

Driver). Khi đăng kí thành công, bộ lọc địa chỉ sẽ gọi lại hàm CALLBACK khi 1 gói tin được gửi ra hay nhận vào hệ thống để xử lí.

### 7.4.2 ***Tóm tắt các bước xây dựng Filter-Hook Driver để bắt gói tin:***

- Khởi tạo Filter-Hook Driver. Cung cấp tên và các thông số cơ bản cho Driver như sau:
  *LoadDriver("IpFilterDriver","System32\\Drivers\\IpFltDrv.sys", null, true)*
- Lấy con trỏ địa chỉ của Ip Filter Driver đã khởi tạo ở bước 1 để khởi tạo và đăng kí hàm CALLBACK.
- Khởi tạo và đăng kí hàm CALLBACK bằng cách gửi con trỏ hàm CALLBACK đã định nghĩa sẵn cho IP Filter Driver.
- Bắt đầu lọc gói tin. Gọi hàm **StartFilter**.
- Khi muốn kết thúc, không lọc gói tin nữa thì ta phải gỡ bỏ thông tin đăng kí khỏi IP Filter Driver. Lúc này, ta chỉ cần đăng kí lại với Driver với con trỏ hàm CALLBACK là Null.

## 7.5 Chi tiết lưu trữ dữ liệu :

### 7.5.1 ***Bảng ForbiddenProxy***

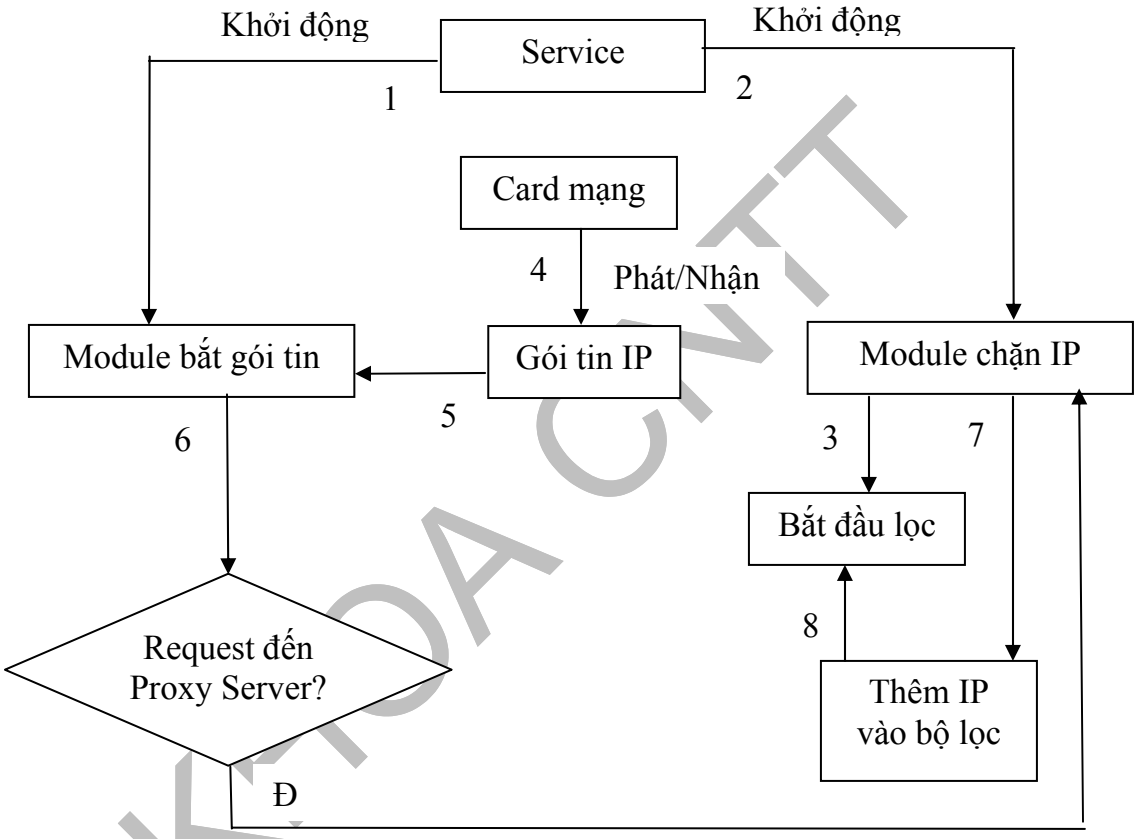
| Tên trường | Kiểu | Chú thích |
|---|---|---|
| ProxyIP | Text | Địa chỉ IP của proxy bị cấm (do service lưu lại được trong quá trình hoạt động) |

### 7.5.2 ***Bảng TrustedProxy:***

| Tên trường | Kiểu | Chú thích |
|---|---|---|

| | | |
|---|---|---|
| ProxyIP | Text | Địa chỉ IP của các Proxy server tin cậy (thường là địa chỉ Proxy Server trong mạng LAN) |

## 7.6 Sơ đồ hoạt động của Module chặn địa chỉ IP :




**Hình 31 Sơ đồ hoạt động của module chặn địa chỉ IP**

## 7.7 Diễn giải mô hình :

Khi khởi động, service sẽ kích hoạt 2 module con là **module bắt gói tin** và **module chặn địa chỉ IP** tương úng. **Module chặn địa chỉ IP** khi được khởi động sẽ truy xuất vào cơ sở dữ liệu và thêm các địa chỉ IP của các Proxy Server bị cấm sẵn vào bộ lọc **IP Filter Driver** và bắt đầu lọc địa chỉ. Khi Card mạng nhận/phát các gói tin, **module bắt gói tin** sẽ nhận các gói tin này và tiến hành phân tích. Module sẽ kiểm tra

xem các gói tin này có phải là gói tin HTTP Request đến Proxy Server hay không. Nếu phải thì địa chỉ IP của Proxy Server sẽ được truyền tiếp cho **Module lọc địa chỉ IP** xử lí. Địa chỉ mới này sẽ được thêm vào bộ lọc địa chỉ và lưu vào cơ sở dữ liệu.

## 7.8 Nhận xét – đánh giá :

### 7.8.1 *Ưu điểm:*

Do yêu cầu đặt ra ban đầu của Module là tìm cách chặn phương pháp vượt Firewall thông qua HTTP Proxy Server, nên chúng em đã cố gắng phát triển module dưới dạng 1 ứng dụng **Mini Firewall**. Trong suốt quá trình nghiên cứu và tìm hiểu, chúng em đã thống nhất chọn **mô hình Service ứng dụng trên Windows** làm cơ sở xây dựng và triển khai Module. Ưu điểm của mô hình này là nó kế thừa được những yêu cầu về **tính an toàn và bảo mật** do chính hệ điều hành qui định. Khi khởi động vào môi trường Windows, các Services hệ thống cũng như của người dùng sẽ lần lượt được nạp và chạy nền trên hệ thống, chỉ duy nhất người quản trị hay chủ Service mới có quyền tắt/mở/xóa service. Module đã nhường hẳn chức năng điều khiển Service cho hệ điều hành, nên Module ứng dụng chỉ tập trung vào hai tính năng chính là **bắt gói tin** và **lọc địa chỉ IP**.

Trong quá trình chạy thử nghiệm, module có thể hoạt động tốt trên các loại CARD mạng, MODEM trên Windows. Do các module con của ứng dụng được viết hoàn toàn dựa trên môi trường Winsock của Windows (bộ thư viện dùng để phát triển ứng dụng mạng TCP/IP trên môi trường Windows), nên bảo đảm tính tương thích rất cao.

Do hỗ trợ tính năng bắt gói tin, Module có thể "phát hiện và học" được các địa chỉ Proxy Server mới (chưa có trong cơ sở dữ liệu). Sau đó lưu lại các địa chỉ này làm cơ sở cho bộ lọc hoạt động

### 7.8.2 ***Khuyết điểm:***

Trong quá trình chạy thử nghiệm, module chặn được gần như hầu hết các địa chỉ HTTP Proxy Server. Tuy nhiên đối với các Proxy Server mới (chưa có trong cơ sở dữ liệu), bộ lọc phải "học" được địa chỉ mới này thì mới ngăn chặn được. Do đó trong phiên làm việc đầu tiên, bộ lọc vẫn chưa chặn được các địa chỉ mới này. Đối với những phiên làm việc sau thì bộ lọc đảm bảo chạy tốt.

Trong quá trình thử nghiệm, việc bộ lọc học được quá nhiều địa chỉ mới và lưu vào cơ sở dữ liệu tốn khá nhiều tài nguyên hệ thống (CPU,RAM) nên Service chạy chậm hẳn (đôi lúc Serive có thể bị treo). Đáng tiếc là đến lúc này chúng em vẫn chưa khắc phục được vấn đề này

Quá trình hoạt động của Service phụ thuộc khá nhiều vào "sự tồn tại" của tập tin cơ sở dữ liệu lưu trữ các Proxy Server. Nên khi tập tin trên không tồn tại hay bị lỗi, tính năng lọc của Service chắc chắn không thể hoạt động chính xác được.

**PHẦN THỨ 4**

# TỔNG KẾT

# Chương 8: KẾT LUẬN

Sau hơn sáu tháng làm luận văn, ít nhiều chúng em cũng đã tìm hiểu tương đối thành công các phương pháp lập trình vượt firewall cũng như những chương trình kèm theo: Http proxy, Web based Proxy, Plug-in chống vượt firewall, service chống vượt firewall. Qua những gì tìm hiểu được, chúng em cảm thấy vẫn còn nhiều điều phải làm để có thể hoàn thiện hơn chương trình cũng như cần có sự hướng dẫn nhiều hơn nữa của các thầy cô, bạn bè…

Kết quả cuối cùng là kết quả của những tháng ngày cố gắng, nỗ lực của bản thân, sự giúp đỡ của gia đình, nhà trường, bạn bè và đặc biệt là sự hướng dẫn tận tình của thầy Đỗ Hoàng Cường để chúng em có thể hoàn tất một cách tốt đẹp luận văn so với những gì đã đặt ra.

Cuối cùng, một lần nữa, chúng em xin cảm ơn tất cả đã giúp đỡ để chúng em có thể hoàn thành tốt khóa luận này. Xin chân thành cảm ơn.

## 8.1 Những kết quả đạt được:

Theo yêu cầu đặt ra ban đầu là ***"Nghiên cứu các phương pháp lập trình vượt firewall. Từ đó làm cơ sở xây dựng các module chống vượt Firewall và bảo mật Web",*** cho đến thời điểm hiện tại luận văn đã đạt được các nội dung sau:

- Phần yêu cầu:
  - Tìm hiểu và triển khai thành công 2 phương pháp: **HTTP Proxy Server** và **Web-based Proxy**.
- Phần mở rộng:
  - Tìm hiểu và triển khai thành công 2 module chống vượt Firewall: **Plugin chống vượt Firewall dành cho trình duyệt Internet Explorer** và **Service chống vượt Firewall trên hệ điều hành Windows.**

Ngoài ra, trong quá trình nghiên cứu và hoàn thành đề tài, chúng em đã tiếp thu thêm được một số kết quả sau:

- Tìm hiểu sâu thêm về các phương pháp lập trình ứng dụng mạng dựa trên bộ thư viện Winsock của Windows.
- Tìm hiểu được phương pháp xây dựng và triển khai Service ứng dụng trên Windows
- Tìm hiểu cách xây dựng và triển khai ứng dụng Plugin cho trình duyệt Internet Explorer.
- Được hiểu được cách xây dựng và phát triển ứng dụng dựa trên môi trường COM (Component Object Model).
- Ngày nay, Internet ngày càng phát triển mạnh mẽ, là nguồn tài nguyên bao la vô tận, nên nhu cầu sử dụng Internet để tìm kiếm thông tin cũng như giao dịch, thương mại là điều tất yếu. Yêu cầu an toàn và bảo mật thông tin (tùy theo mục đích của cá nhân hay doanh nghiệp) đã làm nảy sinh thêm vấn đề khá đau đầu cho các nhà quản trị mạng là: Kiểm soát và quản lý quá trình sử dụng Internet của người dùng. Với việc nghiên cứu và đưa ra được các giải pháp khả thi về yêu cầu mở rộng của đề tài: Xây dựng các module chống vượt Firewall, chúng em thiết nghĩ có thể đóng góp 1 phần vào việc giải quyết vấn đề nan giải trên.

## 8.2 Hướng phát triển :

Trong quá trình nghiên cứu và tìm hiểu về đề tài, chúng em đã thống nhất và đề xuất ra được 3 phương pháp chủ yếu để vượt Firewall: **HTTP Proxy Server**, **Web Based Proxy** và **HTTP Tunneling**. Tất cả 3 phương pháp trên đều được phát triển dựa trên mô hình ứng dụng mạng Client-Server truyền thống. Trong 3 phương pháp nêu trên thì phương pháp thứ 3: **HTTP Tunneling** là phương pháp **cao cấp và khó phát hiện nhất**. Quá trình nghiên cứu và triển khai phương pháp này cũng tốn khá nhiều thời gian và công sức. Mặc dù chúng em đã rất cố gắng triển khai, ý tưởng trên vẫn chưa mang tính khả thi cao và có thể áp dụng được vào thực tế. Sau đây chúng em xin

đề ra 1 số hướng phát triển về sau nhằm mở rộng thêm ý nghĩa khoa học cũng như thực tiễn của đề tài:

- Cải thiện vấn đề tốc độ truy xuất bộ lọc cho module thứ 2: Service chống vượt Firewall.
- Nghiên cứu tiếp phương pháp http tunneling
- Triển khai ứng dụng minh họa cho phương pháp http tunneling
- Hoàn thiện hơn nữa Plug-in và Service để đạt hiệu quả tối ưu
- Triển khai thành công module chống vượt Firewall bằng phương pháp **HTTP Tunneling**
- Triển khai đề tài thành sản phẩm hoàn chỉnh để áp dụng vào thực tiễn.

**PHẦN THỨ 5**

# PHỤ LỤC

## DANH SÁCH CÁC TÀI LIỆU THAM KHÀO

- *Website:*


- **http://www.microsoft.com**
- **http://www.quantrimang.com**
- **http://www.codeproject.com**
- **http://www.sourceforge.net**
- **http://www.experts-exchange.com**
- **http://www.webopedia.com**
- **http://www.nyu.edu**
- **http://www.learnthat.com**
- **http://www.stayinvisible.com**
- **http://www.proxify.com**
- **http://www.silentsurf.net**
- **http://www.adminvietnam.net**
- **http://www.anonimizer.com**
- **http://www.tcpipguide.com**
- **http://www.vnsecurity.net**


- *Danh sách các tài liệu, sách, giáo trình tham khảo*


- Tài liệu điện tử MSDN của Microsoft.
- **Anthony Jones** và **Jim Ohlund**, Network Programming for Microsoft Windows, 1999 (ebooks)
- **O'Reilly**, Learning PHP 5,June-2004
- **Addision Wesley**, The C++ Programming Language,June-97

- **Wrox Press,**Beginning PHP 4,2001
- **Sams Publishing** ,Teach Yourself PHP, MySQL and Apache in 24h,12-2002
- **Addision Wesley**,C/C++ Network Programming I & II,10-2001