ĐẠI HỌC THÁI NGUYÊN
TRƯỜNG ĐẠI HỌC CNTT&TT

Lê Hồng Sơn

GIẢI PHÁP BẢO VỆ WEB SERVER DỰA TRÊN REVERSE PROXY

LUẬN VĂN THẠC SĨ KHOA HỌC MÁY TÍNH

Thái Nguyên - 2012

MỞ ĐẦU

1. Lý do chọn đề tài: Hiện nay, bảo mật thông tin đang đóng một vai trò thiết yếu trong mọi hoạt động liên quan đến việc ứng dụng công nghệ thông tin.

Trong đó Web Server luôn là những vùng đất màu mỡ cho các hacker tìm kiếm các thông tin giá trị hay gây rối vì một mục đích nào đó. Hiểm hoạ có thể là tấn công từ chối dịch vụ, quảng cáo các website có nội dung không lành mạnh, xoá, thay đổi nội dung các file hay phần mềm chứa mã nguy hiểm vv… Các nhà quản trị mạng luôn phải đau đầu, lo lắng tìm các phương pháp để bảo vệ Web server và an toàn thông tin cho toàn bộ hệ thống.

Xuất phát từ những nhu cầu trên, học viên quyết định lựa chọn đề tài "Nghiên cứu giải pháp bảo vệ Web Server dựa trên Reverse Proxy" mong muốn nghiên cứu, đánh giá khả năng bảo vệ Web Server và lựa chọn một ứng dụng để cấu hình và cài đặt thử nghiệm mô hình cụ thể. Ưu điểm của giải pháp này là chi phí để xây dựng hệ thống bảo vệ Web Server là thấp.

2. Đối tượng và phạm vi nghiên cứu: Trong những năm vừa qua, hàng loạt các vụ tấn công vào hệ thống Web Server của các trang mạng xã hội tạo ra những quan tâm rất lớn trong các nhà quản trị hệ thống mạng thông tin.

Sử dụng Reverse Proxy là một cách bảo vệ Web Server. Reverse Proxy đứng giữa một Server và tất cả Client mà Server phải phục vụ, hoạt động như một trạm kiểm soát, các request từ Client bắt buộc phải vào Reverse Proxy. Tại Reverse Proxy sẽ kiểm soát, lọc bỏ các request không hợp lệ và luân chuyển các request hợp lệ đến đích cuối cùng là các Server.

Việc bảo vệ Web Server có nhiều biện pháp khác nhau, do đó đối tượng nghiên cứu của luận văn chỉ tập trung vào việc nghiên cứu giải pháp bảo vệ

Web Server dựa trên Reverse Proxy. Sau đó cấu hình và cài đặt thử nghiệp bằng phần mềm mã nguồn mở NGINX.

3. Hướng nghiên cứu của đề tài: Việc nghiên cứu và đánh giá các vấn đề trong đề tài dựa trên các cơ sở khoa học và phương pháp luận nghiên cứu sau:

Hệ thống lý thuyết tổng quan về bảo mật trên Web Server; SSL; HTTPS; Reverse Proxy.

Nghiêm cứu mô hình triển khai hệ thống mạng bảo mật cho Web Server và cấu hình và cài đặt thử nghiệm Reverse Proxy bằng phần mềm NGINX.

4. Phương pháp nghiên cứu: Thu nhập hệ thống lý thuyết tổng quan về bảo mật trên web server; SSL; HTTPS; Reverse Proxy. Tham khảo các tài liệu liên quan ở trong nước, nước ngoài và trên Internet; sử dụng phương pháp phân tích, liệt kê, thực nghiệm,…

5. Ý nghĩa khoa học của đề tài: Nghiên cứu tìm hiểu giải pháp bảo vệ Web Server dựa trên Reverse Proxy, cấu hình và cài đặt thử nghiệp bằng phần mềm mã nguồn mở NGINX.

Vì điều kiện thời gian có giới hạn và năng lực của bản thân, mặc dù đã cố gắng nhưng đề tài chắc có thể chưa đi sâu phân tích hết các khía cạnh, chi tiết có liên quan. Kính mong Thầy hướng dẫn và Hội đồng bảo vệ luận văn tốt nghiệp cho ý kiến đóng góp thêm để đề tài được hoàn thiện hơn.

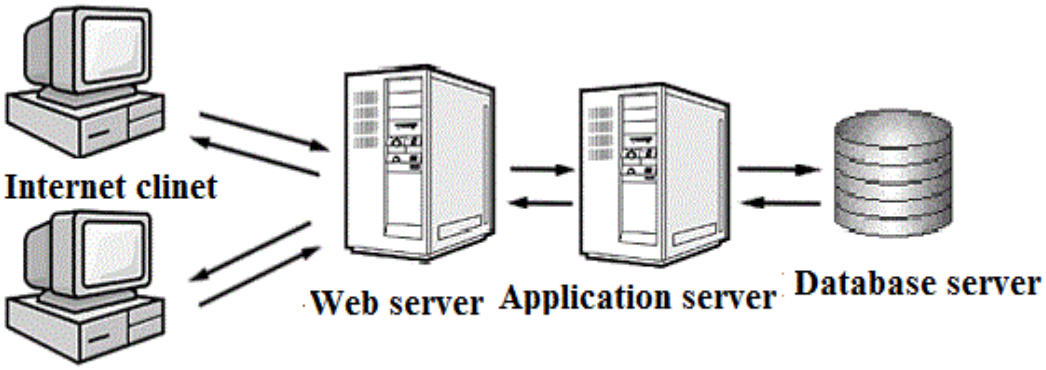
CHƯƠNG 1:
TỔNG QUAN CÁC GIẢI PHÁP BẢO VỆ WEB SERVER

1.1.Tổng quan Web Server

Web Server là máy chủ có dung lượng lớn, tốc độ cao, được dùng để lưu trữ thông tin như một ngân hàng dữ liệu, chứa những website đã được thiết kế cùng với những thông tin liên quan khác (các mã Script, các chương trình, và các file Multimedia).

Web Server gửi đến Client những trang Web thông qua môi trường Internet qua giao thức HTTP, HTTPS; giao thức được thiết kế để gửi các file đến Web Browser và các giao thức khác.

Khi máy tính kết nối đến một Web Server và gửi đến yêu cầu truy cập các thông tin từ một trang Web nào đó, Web Server sẽ nhận yêu cầu và gửi lại những thông tin yêu cầu. Giống các phần mềm khác Web Server cũng chỉ là một ứng dụng phần mềm. Nó được cài đặt, chạy trên máy tính dùng làm Web Server, nhờ có chương trình này mà người sử dụng có thể truy cập đến các thông tin của trang Web từ một máy tính khác ở trên mạng Internet, Intranet.



Hình 1.1 Mô hình hoạt động Web Server

Web Server còn có thể được tích hợp với Database hay điều khiển việc kết nối vào Database để có thể truy cập và kết xuất thông tin từ Database lên các trang Web và truyền tải chúng đến người dùng. Web Server phải hoạt

động liên tục 24/24 giờ, 7 ngày một tuần và 365 ngày một năm, để phục vụ cho việc cung cấp thông tin trực tuyến.

Có nhiều Web Server khác nhau, việc lựa chọn một web server phù hợp sẽ dựa trên các tiêu chí đánh giá. Khả năng làm việc với hệ điều hành, các ứng dụng khác, thiết lập các chương trình ứng dụng phía server, bảo mật dữ liệu, xuất bản trang web, các công cụ hỗ trợ khi xây dựng các trang web.

Hiện nay, có 2 loại web server thông dụng nhất là : Internet Information Services (IIS), Apache Web Server.

1.1.1. Internet Information Services (IIS)

Internet Information Services (IIS) là một dịch vụ tùy chọn của Windows Server cung cấp các tính năng về Web site.

Giải pháp phổ biến nhất của Microsoft cho một web site là chạy IIS trên nền Windows Server. IIS là dịch vụ thông tin Internet do Microsoft phát triển, sản phẩm này được tích hợp cùng với hệ điều hành Windows. Phiên bản mới nhất hiện nay là IIS 7.0 được chạy trên hệ điều hành Windows Server 2003, 2008... Phiên bản này được Microsoft thiết kế lại dưới dạng module, vừa kế thừa ưu điểm của những phiên bản trước,vừa tăng cường tính bảo mật và ổn định. Những điểm đáng chú ý trong IIS 7.0 bao gồm:

IIS 7.0 cung cấp 2 công cụ quản trị, một dưới dạng đồ họa và một dưới dạng dòng lệnh. Những công cụ quản trị này cho phép bạn:

- Quản lý tập trung IIS và ASP.NET;
- Xem thông tin, chẩn đoán, trong đó bao gồm các thông tin real-time;
- Thay đổi quyền trên các đối tượng site và ứng dụng;
- Ủy quyền cấu hình các đối tượng site và ứng dụng cho các thành viên không có quyền quản trị;

Hình 1.2 Phần mềm IIS7

Thay đổi cách thức lưu trữ thông tin cấu hình IIS 7.0 và ASP.NET vào một vị trí, từ đó cho phép:

- Cấu hình IIS và ASP.NET với một định dạng thống nhất

- Dễ dàng sao chép các file cấu hình và nội dung của site hoặc ứng dụng đến một máy tính khác

Dễ dàng chẩn đoán và khắc phục sự cố nhờ vào thông tin real-time và hệ thống file log ở mức độ chi tiết

IIS 7.0 được thiết kế dưới dạng module, cho phép bổ sung cũng như loại bỏ các thành phần từ Web Server khi cần.

Khả năng tương thích cao đối với các ứng dụng đã triển khai trong các phiên bản IIS trước. Khi đó triển khai IIS 7.0 có thể chạy các ứng dụng ASP hoặc ASP.NET 2.0 đã được xây dựng từ trước mà không cần phải thay đổi mã nguồn.

Trong IIS bao gồm nhiều dịch vụ dịch vụ như: dịch vụ Web Server, dịch vụ FTP Server … Ở đây chỉ đề cập đến dịch vụ Web Server. IIS Web Server

đáp ứng mọi yêu cầu chủ yếu của một Web Server như: độ tin cậy, hiệu năng, khả năng theo dõi giám sát, tính bảo mật và tính khả thi trong việc phát triển các dịch vụ ứng dụng. Tất cả các cải tiến này là kết quả là sự kết hợp chặt chẽ cùng với các tính năng mới được cung cấp trong hệ điều hành Windows.

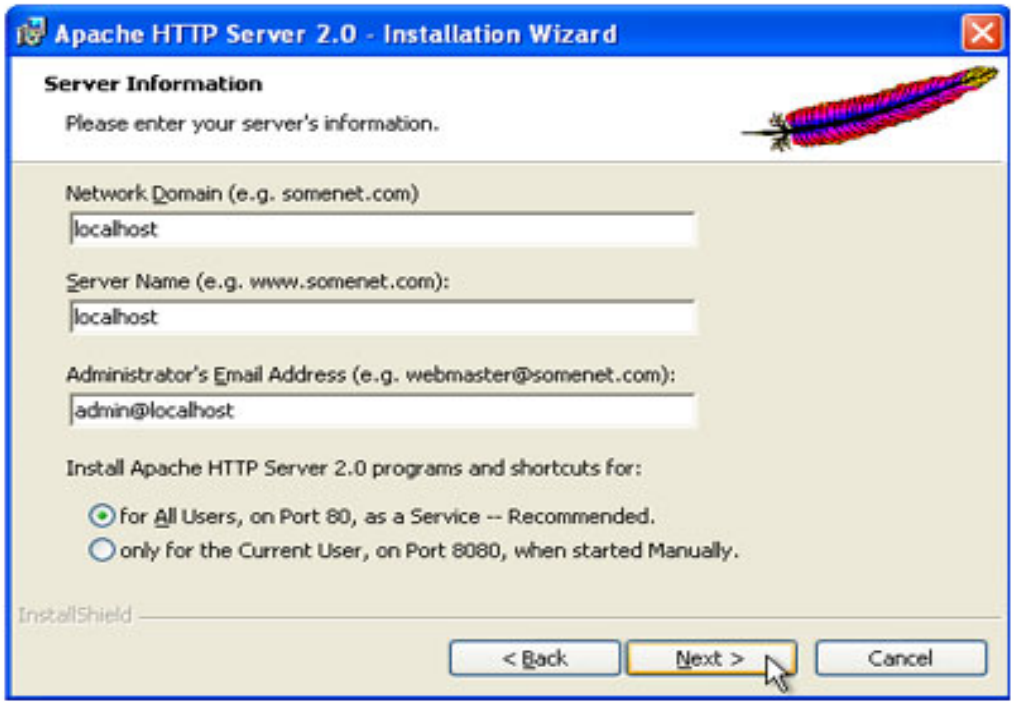
Người dùng có thể triển khai liên tục các ứng dụng mạng lên các Server với nội dung mới nhất. IIS hỗ trợ đầy đủ các hệ thống ngôn ngữ lập trình như Visual Basic, Visual Basic Script, J script^{MT} được phát triển bởi Microsoft và Java Conponent, ứng dụng CGI dành cho các ngôn ngữ lập trình Web cơ sở, ISAPI mở rộng và các bộ lọc.

1.1.2. Apache Web Server

Apache Web Server được xem như một sự nỗ lực rất lớn trong việc phát triển và duy trì một Web Server mã nguồn mở cho các hệ điều hành, bao gồm Unix, Linux và Windows. Đây là một Web Server hội tụ tất cả các tính năng: bảo mật, hiệu suất, mở rộng và phát triển cung cấp các dịch vụ Web được đồng bộ trong các chuẩn Web hiện hành.

Các đặc điểm nổi bật của Apache:

- Apache có thể chạy kết hợp giữa chế độ đa xử lý và chế độ đa chỉ lệnh.
- Hỗ trợ nhiều giao thức: Apache được phát triển để có thể phục vụ trên nhiều giao thức khác nhau.
- Ngày càng hỗ trợ tốt hơn cho các hệ điều hành khác như: Linux, OS và Windows.
- Ngày càng phát triển và hoàn thiện các API (Application Program Interface).
- Hỗ trợ IPv6.
- Hỗ trợ nhiều modul dùng để lọc các dòng dữ liệu đến hoặc đi từ server.



Hình 1.3 Phần mềm apache

- Hỗ trợ nhiều ngôn ngữ hiển thị các thông báo lỗi.

Ngày càng đơn giản và dễ dàng thiết lập các tham số cho Web Server qua các file cấu hình.

1.2. Một số phương thức tấn công web server

1.2.1. Authentication attacks

Authentication đóng một vai trò rất quan trọng trong việc đảm bảo tính an ninh của một web application. Khi một user cung cấp login name và password để xác thực tài khoản của mình, web application cấp quyền truy xuất cho user dựa vào login name mà user nhập vào đã được lưu trong cơ sở dữ liệu. HTTP có một số phương thức xác thực:

- Basic
- Digest
- Form-based.
- NTLM.

- Negotiate.
- Client-side.
- Microsoft Passport.

Kiểu tấn công này không dựa vào lỗ hổng an ninh trên hệ điều hành và phần mềm của server. Nó phụ thuộc vào mức độ an ninh và phức tạp của password được lưu trữ và mức độ khó khăn để cho attacker có thể tiếp cận được server. Khi thực hiện tấn công này, hacker có thể vượt rào xác thực và vào hệ thống với quyền truy xuất mà mình mong muốn. Với quyền đăng nhập cao nhất admin, hacker có thể toàn quyền điều khiển hệ thống web bị tấn công.

Giải pháp tốt hơn cho vấn đề này là sử dụng một vài hình thức "**multi-factor authentication**" (chứng thực sử dụng nhiều yếu tố).

Vấn đề ở đây là sức mạnh tính toán của các máy tính ngày này ngày càng tăng. Chúng có khả năng xử lý một lượng lớn dữ liệu chỉ trong một khoảng thời gian ngắn. Một "password" chỉ là một chuỗi các ký tự (có trên bàn phím) mà một người cần ghi nhớ và cung cấp cho máy tính khi cần thiết (như để đăng nhập vào máy tính, truy cập tài nguyên trên mạng…).

Thật không may, các mật khẩu mà quá phức tạp để ghi nhớ đối với con người thì lại dễ dàng bị dò ra bởi các công cụ "pasword cracking" trong một khoảng thời gian ngắn đến kinh ngạc. Các kiểu tấn công như "dictionary attack", "brute fore attack" và "hybrid attack" thường được sử dụng để đoán và bẻ khóa mật khẩu.

Phương thức bảo vệ duy nhất chống lại những "threat" như vậy là tạo ra các mật khẩu mạnh "**strong password**" (độ dài của mật khẩu thường từ 8 ký tự trở lên, trong đó bao gồm cả chữ cái in thường/in hoa, chữ số, ký tự đặc biệt) và sử dụng thêm các yếu tố khác (vân tay, smart card, võng mạc mắt,…) cho việc chứng thực.

Nhưng ngay cả khi người ta có thể nhớ được các "strong password" (tất nhiên độ phức tạp của "password" này cần ở mức vừa phải) như dài từ 12 đến 16 ký tự, thì vẫn còn các vấn đề khác mà các hệ thống chứng thực chỉ dựa vào "password" phải đối mặt.

1.2.2. HTTP Response Splitting

Lỗi HTTP Response Splitting tấn công vào ứng dụng web và diễn ra khi nó không thể xử lý đúng các thông tin đầu vào người dùng nhập.

Kẻ tấn công từ xa có thể gửi một yêu cầu HTTP đặc biệt làm cho máy chủ web định dạng yêu cầu nhầm tưởng rằng nó chứa 2 yêu cầu HTTP chứ không phải một. Chỉ yêu cầu thứ nhất được xử lý bởi người sử dụng. HTTP Response Splitting cho phép tiến hành một lượng lớn các cuộc tấn công kiểu như web cache positioning, deface, "cross-user defacement", chặn và ăn cắp thông tin người dùng và Cross site Scripting.

1.2.3. File Inclusion Attacks

Khi một trang web sử dụng các lệnh include, require,… để gọi đến một file khác, sơ ý để người dùng có thể thay đổi file cần gọi đến. Như vậy Website đó đang đứng trước nguy cơ bị tấn công File Inclusion. Tùy vào mức độ bảo mật của Server, hacker có thể include đến file trên Server đó hoặc include đến file trên Server khác (remote include). Với từng mức độ hacker có thể có nhiều cách để up shell. Nếu server kém bảo mật thì kẻ tấn công có thể đọc được file của toàn bộ hệ thống, file của các Website khác trên cùng máy chủ đó.

Lỗi này dùng để tấn công local kiểu tấn công máy chủ, website qua một site bị lỗi trên máy chủ. Nếu có quyền ghi, tất cả các file có thể bị thay đổi: deface trang chủ, chèn mã độc để thu thập thông tin đăng nhập, ẩn dấu backdoor để lần sau vào tiếp,…Có thể lấy thông tin truy nhập cơ sở dữ liệu