HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

CƠ SỞ TẠI TPHCM

KHOA CÔNG NGHỆ THÔNG TIN II

BỘ MÔN KHOA HỌC PHÁP LÝ SỐ

BÁO CÁO CUỐI KÌ
ĐỀ TÀI: INVESTIGATING WEB ATTACKS

Giảng viên hướng dẫn : ThS Nguyễn Hoàng Thành

Họ và tên sinh viên : Dương Minh Phong – N20DCAT040
Phan Tiến Sĩ - N20DCAT048
Nguyễn Tuấn Kiệt – N20DCAT028
Phan Công Thăng – N20DCAT057

Lớp :D20CQAT01-N

TPHCM – tháng 5 năm 2024

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

CƠ SỞ TẠI TPHCM

KHOA CÔNG NGHỆ THÔNG TIN II

BỘ MÔN KIỂM THỬ XÂM NHẬP

BÁO CÁO CUỐI KÌ
ĐỀ TÀI: INVESTIGATING WEB ATTACKS

Giảng viên hướng dẫn : **ThS Phan Nghĩa Hiệp**

Họ và tên sinh viên : **Dương Minh Phong – N20DCAT040**
Phan Tiến Sĩ - N20DCAT048
Nguyễn Tuấn Kiệt – N20DCAT028
Phan Công Thăng – N20DCAT057

Lớp :**D20CQAT01-N**

TPHCM – tháng 5 năm 2024

Nhận xét của GVHD

.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....
.....

Lời cảm ơn

Nhóm chúng em xin cảm ơn thầy Nguyễn Hoàng Thành! Trong quá trình học tập, nhờ có sự giúp đỡ và hỗ trợ từ phía thầy nhóm chúng em được hiểu biết thêm nhiều kiến thức, nhận biết ra nhiều cách điều tra khi bị tội phạm mạng tấn công trong thời đại sử dụng cộng nghệ có kết nối mạng thời nay. Không chỉ từ những kiến thức đó giúp chúng em có thêm kỹ năng để tự phòng chống và bảo vệ hệ thống thông tin của chính mình mà phần nào khi ra trường, khi đi làm chúng em có thêm kinh nghiệm để phát triển và tìm kiếm được
nhiều cơ hội.

Đặc biệt, trong quá trình soạn và làm báo cáo đề án cuối kỳ, thầy còn tận tình giúp đỡ chúng em qua việc trả lời các thắc mắc. Và không chỉ riêng chúng em thầy còn hỗ trợ và giúp đỡ các nhóm khác trong lớp được thầy đảm nhiệm và phân công.

Một lần nữa, nhóm chúng em xin được cảm ơn thầy Nguyễn Hoàng Thành giảng viên bô môn Khoa học Pháp lý số.

Tp. Hồ Chí Minh, tháng 05 năm 2024
Nhóm 11

Mục lục

Chương I: Cơ sở lí thuyết

1. Đặt vấn đề

1.1 Web và website là gì ?

Trong thời đại công nghệ số hiện nay, việc truy cập từ xa vào các hệ thống công nghệ thông tin ngày càng phổ biến. Không chỉ là các ứng dụng(Application) mà là còn các Web. Song song với việc phổ biến của các ứng dụng thì việc đảm bảo an toàn cho chúng cũng là việc nên đặt lên hàng đầu. Vậy Web là gì?

Web là tên thường gọi của World Wide Web (mạng toàn cầu), một tập hợp con của Internet bao gồm các trang có thể được truy cập bằng trình duyệt Web.

Web chỉ là một trong những cách chia sẻ thông tin qua Internet bên cạnh những thứ khác bao gồm email, nhắn tin tức thời và Giao thức truyền tệp (FTP).

Trang web, trong tiếng Anh là "web page", là một phần của website. Một website thông thường sẽ bao gồm nhiều web page hoặc tối thiểu là một web page.

Tuy nhiên trong thực tế thì phần lớn người Việt Nam vẫn sẽ ngầm hiểu rằng "trang web" = "website", mặc dù có chút không đúng về mặt định nghĩa nhưng trong giao tiếp hàng ngày, không phải các văn bản quan trọng thì cách sử dụng này vẫn được chấp nhận rộng rãi.[1]

Website thường chứa nhiều webpage hay còn thường được gọi là trang con. Tất cả được lưu trữ dưới định dạng html hoặc xhtml (Extensible HyperText Markup Language - mở rộng của html). Chúng sẽ được lưu trên các máy chủ (web server). Khi người dùng muốn truy cập các thông tin từ website cần sử dụng các trình duyệt web để truy cập vào địa chỉ của website, đọc các file lưu trữ dưới định dạng html hoặc xhtml và hiển thị dưới dạng trực quan để dễ dàng tiếp nhận nội dung, thao tác.

Hiểu một cách ngắn gọn thì Web là mạng.[1]

Còn Site là địa điểm. Ví dụ worksite có nghĩa là nơi làm việc, chỉ một địa điểm, địa chỉ cụ thể.

Như vậy Website = Web + Site, tức một địa chỉ cụ thể trong mạng toàn cầu hay còn được gọi là trang mạng.

Điều đó có nghĩa là website phải đảm bảo được các yếu tố: Nằm trong mạng toàn cầu, có thể truy cập bằng các trình duyệt web, sử dụng giao thức HTTP để truyền dữ liệu và chia sẻ thông tin và quan trọng là phải có một địa chỉ cụ thể.

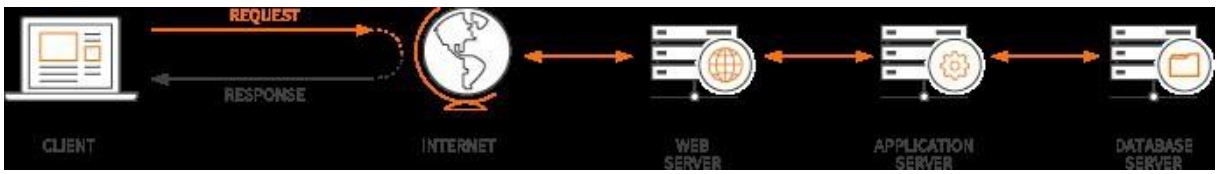
Website thường chứa các nội dung văn bản, hình ảnh, video và rất nhiều định dạng nội dung khác, được lưu trữ trên máy chủ.

Hình 1: Các thành phần của Website

1.2 Web app là gì?

Web Application (Web App) hay Ứng dụng Web là một loại chương trình máy tính thường chạy với sự hỗ trợ của trình duyệt web và công nghệ web để thực hiện các tác vụ khác nhau trên internet. Web Application thường được lưu trữ trên một máy chủ từ xa và người dùng có thể truy cập nó thông qua việc sử dụng Phần mềm được gọi là trình duyệt web.[2]

Các Web Application có thể được thiết kế cho nhiều mục đích sử dụng khác nhau và có thể được sử dụng bởi bất kỳ ai, một tổ chức hoặc một cá nhân. Không giống như các ứng dụng máy tính để bàn, các Web Application có thể được truy cập ở mọi nơi bằng trình duyệt web như Microsoft Explorer, Google Chrome hoặc Apple Safari.



Hình 2: Cách hoạt động của Webapp

- Luồng hoạt động của web app

Người dùng kích hoạt request tới web server qua Internet, thông qua trình duyệt web hoặc giao diện người dùng của ứng dụng.

• Web server chuyển tiếp request này đến web application server thích hợp.
• Máy chủ ứng dụng Web (web application server) thực hiện nhiệm vụ được yêu cầu - chẳng hạn như truy vấn cơ sở dữ liệu hoặc xử lý dữ liệu - sau đó tạo ra các kết quả của dữ liệu được yêu cầu.
• Máy chủ ứng dụng web gửi kết quả đến máy chủ web với thông tin được yêu cầu hoặc dữ liệu đã được xử lý.
• Máy chủ web phản hồi response lại cho khách hàng các thông tin được yêu cầu sau đó xuất hiện trên màn hình của người dùng.[2]

- **Ưu điểm của web app**

+ Web Application không giới hạn cho một nền tảng cụ thể. Chúng có thể được xây dựng cho tất cả các nền tảng miễn là chúng có thể chạy trong trình duyệt web, dễ dàng tương thích ứng với iOS, Android hoặc Windows[2]

+ Ứng dụng web không yêu cầu nhiều thời gian hoặc nhân lực để xây dựng, không cần thử nghiệm trong mỗi hệ điều hành khác nhau, do đó nó có chi phí đầu tư thấp hơn nhiều so với các loại phát triển ứng dụng khác.

+ Web App sử dụng các ngôn ngữ mã hóa phổ biến trên nhiều nền tảng, do đó việc xây dựng và bảo trì tương đối dễ dàng

+ Cập nhật tự động, người dùng luôn thấy phiên bản cập nhật nhất khi họ mở ứng dụng web.

- **Nhược điểm của web app**

+ Web Application phải được truy cập thông qua trình duyệt web; do đó, nếu không có kết nối internet, người dùng sẽ không thể truy cập bất kỳ ứng dụng web nào.

+ Ứng dụng web có ít chức năng hơn các loại ứng dụng khác, chúng không có quyền truy cập vào các tính năng và phần cứng của thiết bị.[2]

+Các Web App có UX kém, nên việc cung cấp trải nghiệm liền mạch cho người dùng trở nên khó khăn hơn.

+ Web Application được liên kết trực tiếp với trình duyệt web. Điều đó có nghĩa là, nếu trang web xảy ra lỗi thì rất có thể ứng dụng cũng sẽ bị lỗi.

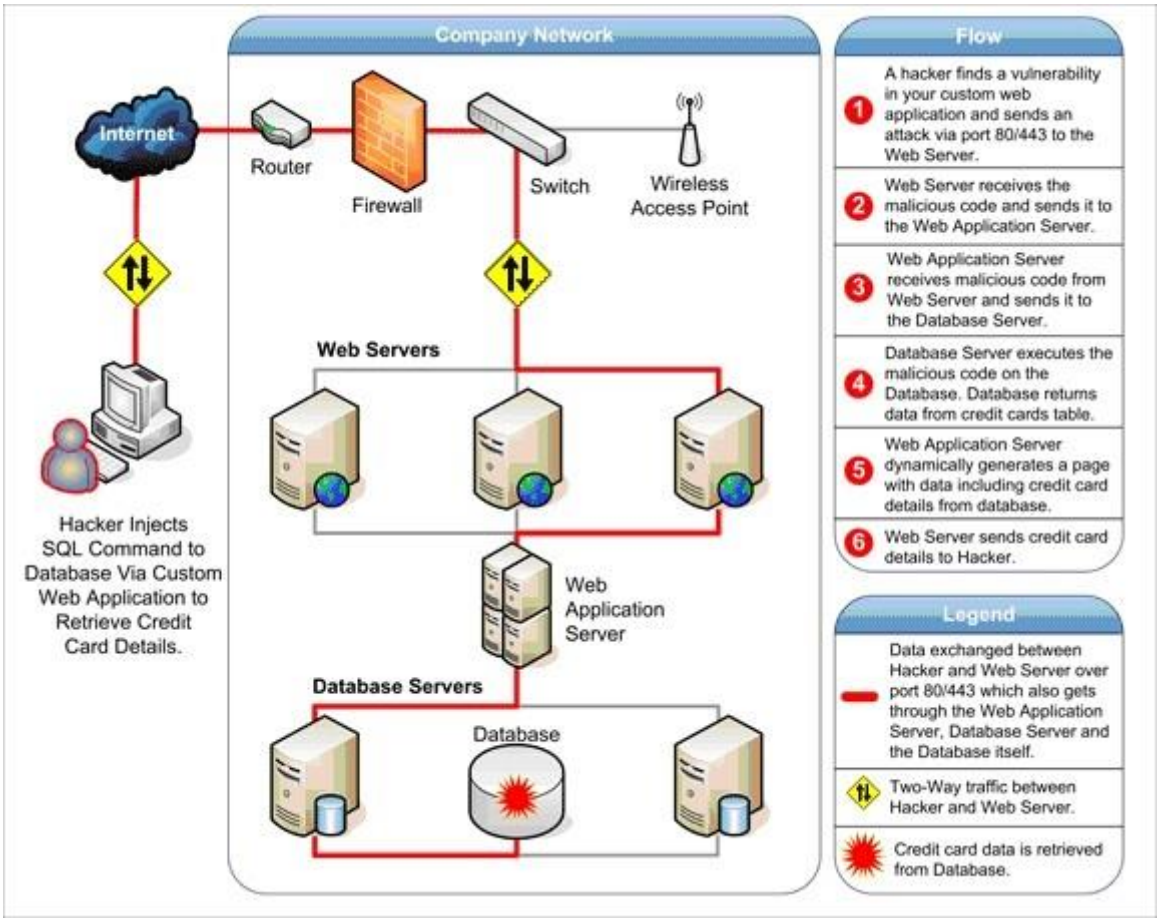
+Web App thiếu tính năng của hệ thống kiểm soát chất lượng. Do đó, tính an toàn và bảo mật đều không cao

2. Investigating web attacks là gì?

Bây giờ chúng ta hãy xem xét các kiểu tấn công vào các ứng dụng web. Bất chấp những ưu điểm của chúng, các ứng dụng web gây ra một số lo ngại về bảo mật bắt nguồn từ việc mã hóa không đúng cách. Các điểm yếu hoặc lỗ hổng nghiêm trọng cho phép bọn tội phạm truy cập trực tiếp và công khai vào cơ sở dữ liệu để đánh cắp dữ liệu nhạy cảm - đây được gọi là một cuộc tấn công ứng dụng web. Nhiều cơ sở dữ liệu trong số này chứa thông tin có giá trị (ví dụ: dữ liệu cá nhân và chi tiết tài chính) khiến chúng trở thành mục tiêu tấn công thường xuyên. Mặc dù những hành động phá hoại như vậy (thường được thực hiện bởi những kẻ được gọi là kẻ lừa đảo tập lệnh) như làm xấu các trang web của công ty vẫn diễn ra phổ biến, nhưng ngày nay những kẻ tấn công thích giành quyền truy cập vào dữ liệu nhạy cảm nằm trên máy chủ cơ sở dữ liệu vì lợi nhuận lớn khi bán kết quả của vi phạm dữ liệu. Trong khuôn khổ mô tả ở trên, có thể dễ dàng thấy tội phạm có thể nhanh chóng truy cập dữ liệu nằm trên cơ sở dữ liệu thông qua một liều lượng sáng tạo và cùng với sự may rủi, sơ suất hoặc lỗi của con người, dẫn đến lỗ hổng trong các ứng dụng web.

Như đã nêu, các trang web phụ thuộc vào cơ sở dữ liệu để cung cấp thông tin cần thiết cho khách truy cập. Nếu các ứng dụng web không an toàn, tức là dễ bị tấn công bởi ít nhất một trong các hình thức kỹ thuật tấn công khác nhau, thì toàn bộ cơ sở dữ liệu thông tin nhạy cảm của bạn có nguy cơ bị tấn công ứng dụng web. Các kiểu tấn công SQL Injection, nhắm mục tiêu trực tiếp vào cơ sở dữ liệu, vẫn là loại lỗ hổng phổ biến nhất và nguy hiểm nhất. Những kẻ tấn công khác có thể tiêm mã độc hại bằng cách sử dụng đầu vào của người dùng của các ứng dụng web dễ bị tấn công để lừa người dùng và chuyển hướng họ đến các trang lừa đảo. Loại tấn công này được gọi là Cross-Site Scripting (tấn công XSS) và có thể được sử dụng ngay cả khi bản thân các máy chủ web và công cụ cơ sở dữ liệu không chứa lỗ hổng bảo mật. Nó thường được sử dụng kết hợp với các vectơ tấn công khác như các cuộc tấn công kỹ thuật xã hội. Có nhiều kiểu tấn công phổ biến khác như duyệt qua thư mục, bao gồm tệp cục bộ, v.v.

Nghiên cứu gần đây cho thấy 75% các cuộc tấn công mạng được thực hiện ở cấp độ ứng dụng web.



Hình 3: Mô hình tấn công SQL Injection

Các ứng dụng web thường có quyền truy cập trực tiếp vào dữ liệu phụ trợ như cơ sở dữ liệu khách hàng và do đó, kiểm soát dữ liệu có giá trị và khó bảo mật hơn nhiều. Những người không có quyền truy cập sẽ có một số dạng tập lệnh cho phép thu thập và truyền dữ liệu. Nếu kẻ tấn công nhận ra điểm yếu trong một tập lệnh như vậy, chúng có thể dễ dàng định tuyến lại lưu lượng truy cập không chủ ý đến một vị trí khác và cung cấp thông tin cá nhân một cách bất hợp pháp.

Do đó, các ứng dụng web là một cổng vào cơ sở dữ liệu, đặc biệt là các ứng dụng tùy chỉnh không được phát triển với các phương pháp bảo mật tốt nhất và không trải qua các cuộc kiểm tra bảo mật thường xuyên. Nói chung, bạn cần trả lời câu hỏi: "Phần nào của trang web mà chúng tôi cho là an toàn, đang mở ra cho một cuộc tấn công ứng dụng web?" và “chúng ta có thể ném dữ liệu nào vào một ứng dụng để khiến nó thực hiện những điều mà nó không nên làm?”.

- ⇨ Tóm lại việc điều tra các trang web bị tấn công là dựa vào các chứng cứ mà thủ phạm để lại để biết được rằng hacker đã làm gì đối với dữ liệu của công ty. Những dữ liệu nào đã bị đánh cắp? những dữ liệu nào đã bị xoá? Thủ phạm

xâm nhập vào hệ thống bằng con đường nào? Đây là một công việc quan trọng bắt buộc phải làm để đảm bảo sự toàn vẹn thông tin của doanh nghiệp cũng như nâng cao sức phòng thủ của trang web cũng như đội phòng thủ trong doanh nghiệp.[3]

3. Các dạng tấn công web

Các loại tấn công Web khác nhau trong thực tế như sau:

• Tấn công kịch bản chéo trang (XSS)

Tấn công Cross Site Scripting nghĩa là gửi và chèn lệnh và script độc hại, những mã độc này thường được viết với ngôn ngữ lập trình phía client như Javascript, HTML, VBScript, Flash… Tuy nhiên, cách tấn công này thông thường sử dụng Javascript và HTML. Cách tấn công này có thể được thực hiện theo nhiều cách khác nhau, phụ thuộc vào loại tấn công XSS, những mã độc có thể được phản chiếu trên trình duyệt của nạn nhân hoặc được lưu trữ trong cơ sở dữ liệu và được chạy mỗi khi người dùng gọi chức năng thích hợp. Nguyên nhân chính của loại tấn công này là xác thực đầu vào dữ liệu người dùng không phù hợp, dữ liệu độc hại từ đầu vào có thể xâm nhập vào dữ liệu đầu ra. Mã độc có thể nhập một script và được chèn vào mã nguồn của website. Khi đó trình duyệt không thể biết mã thực thi có phải độc hại hay không. Do đó mã độc hại có thể đang được thực thi trên trình duyệt của nạn nhân hoặc bất kỳ hình thức giả nào đang được hiển thị cho người sử dụng. Có một số hình thức tấn công XSS có thể xảy ra.

• Giả mạo yêu cầu trên nhiều trang (CSRF)

Cross-site request forgery (CSRF) là một loại tấn công mạng, trong đó kẻ tấn công lừa người dùng thực hiện các hành động không mong muốn trên một trang web mà họ đã đăng nhập bằng cách sử dụng quyền truy cập của người dùng đó. Thường thì người dùng sẽ không nhận ra rằng họ đang thực hiện các hành động này, vì vậy tấn công CSRF được coi là một trong những mối đe dọa nghiêm trọng đối với an ninh của các trang web. Các tấn công CSRF thường xảy ra khi kẻ tấn công tạo ra một yêu cầu giả mạo hoặc chèn một mã độc vào một trang web khác, khiến cho khi người dùng truy cập vào trang web đó và đăng nhập, yêu cầu giả mạo được thực hiện và các hành động không mong muốn được thực hiện mà không được sự cho phép của người dùng.

• SQL injection

SQL Injection là một kỹ thuật lợi dụng những lỗ hổng về câu truy vấn của các ứng dụng. Được thực hiện bằng cách chèn thêm một đoạn SQL để làm sai lệnh đi câu truy

vấn ban đầu, từ đó có thể khai thác dữ liệu từ database. SQL injection có thể cho phép những kẻ tấn công thực hiện các thao tác như một người quản trị web, trên cơ sở dữ liệu của ứng dụng.
Ví dụ, trong form đăng nhập, người dùng nhập dữ liệu, trong trường tìm kiếm người dùng nhập văn bản tìm kiếm, trong biểu mẫu lưu dữ liệu, người dùng nhập dữ liệu cần lưu. Tất cả các dữ liệu được chỉ định này đều đi vào cơ sở dữ liệu.

Thay vì nhập dữ liệu đúng, kẻ tấn công lợi dụng lỗ hổng để insert và thực thi các câu lệnh SQL bất hợp pháp để lấy dữ liệu của người dùng… SQL Injection được thực hiện với ngôn ngữ lập trình SQL. SQL (Structured Query Language) được sử dụng để quản lý dữ liệu được lưu trữ trong toàn bộ cơ sở dữ liệu.

• Code injection
Tấn công code injection tương tự như tấn công tiêm SQL. Trong cuộc tấn công này, khi người dùng gửi bất kỳ ứng dụng nào tới máy chủ, kẻ tấn công sẽ hack ứng dụng và thêm mã độc, chẳng hạn như lệnh shell hoặc tập lệnh PHP.
Khi máy chủ nhận được yêu cầu, nó sẽ thực thi ứng dụng đó. Mục tiêu chính của cuộc tấn công này là bỏ qua hoặc sửa đổi chương trình gốc để thực thi mã tùy ý và giành quyền truy cập vào các trang Web hoặc cơ sở dữ liệu bị hạn chế, bao gồm cả những trang có thông tin cá nhân như số thẻ tín dụng và mật khẩu

• Giả mạo tham số
Giả mạo tham số là một kiểu tấn công Web xảy ra khi kẻ tấn công thay đổi hoặc sửa đổi các tham số của một URL là một chuỗi định danh cho biết nơi đặt tài nguyên và cơ chế cần thiết để lấy tài nguyên đó. Việc giả mạo tham số lợi dụng lập trình viên dựa vào các trường ẩn hoặc cố định, chẳng hạn như thẻ ẩn trong biểu mẫu hoặc tham số trong URL, làm biện pháp bảo mật duy nhất để bảo vệ dữ liệu của người dùng. Kẻ tấn công rất dễ dàng sửa đổi các tham số này.

• Cookie poisoning
Các ứng dụng web sử dụng cookie để lưu trữ thông tin như ID người dùng, mật khẩu, số tài khoản và thời gian đăng nhập, tất cả đều có trên máy cục bộ của người dùng. Trong một cuộc tấn công đầu độc cookie, kẻ tấn công sửa đổi nội dung của cookie để đánh cắp thông tin cá nhân về người dùng hoặc lừa đảo các trang Web

• Buffer overflow
Buffer là vùng lưu trữ dữ liệu tạm thời, có dung lượng hạn chế. Nếu một chương trình lưu trữ nhiều dữ liệu trong bộ đệm hơn mức có thể xử lý, bộ đệm sẽ tràn và tràn dữ

liệu vào một bộ đệm hoàn toàn khác, ghi đè hoặc làm hỏng bộ đệm. dữ liệu hiện có trong bộ đệm đó. Trong các cuộc tấn công như vậy, dữ liệu bổ sung có thể chứa mã độc.
Cuộc tấn công này có thể thay đổi dữ liệu, làm hỏng tệp hoặc tiết lộ thông tin cá nhân. Để thực hiện tràn bộ đệm kẻ tấn công sẽ cố gắng làm tràn các máy chủ phụ trợ với các yêu cầu vượt quá. Sau đó họ gửi được chế tạo đặc biệt cho phép kẻ tấn công kiểm soát các ứng dụng. Cả ứng dụng Web và các sản phẩm máy chủ, hoạt động như các tính năng tĩnh hoặc động của trang web, dễ xảy ra lỗi tràn bộ đệm. lỗi tràn được tìm thấy trong các sản phẩm máy chủ thường được biết đến.

• Rò rỉ cookie
Cookie snooping là khi kẻ tấn công đánh cắp cookie của nạn nhân, có thể sử dụng proxy cục bộ và sử dụng chúng để đăng nhập với tư cách là nạn nhân. Việc sử dụng cookie được mã hóa mạnh và nhúng địa chỉ IP nguồn vào cookie có thể ngăn chặn điều này. Cơ chế cookie có thể được tích hợp hoàn toàn với chức năng SSL để tăng cường bảo mật.

• Tấn công giao thức DMZ
Hầu hết các môi trường ứng dụng Web bao gồm các giao thức như DNS và FTP. Các giao thức này có các lỗ hổng cố hữu thường xuyên bị khai thác để có quyền truy cập vào các tài nguyên ứng dụng quan trọng khác.
DMZ (khu vực phi quân sự) là vùng mạng bán tin cậy tách biệt Internet không tin cậy khỏi mạng nội bộ đáng tin cậy của công ty. Để tăng cường an ninh của DMZ và giảm rủi ro, hầu hết các công ty đều hạn chế các giao thức được phép truyền qua DMZ của họ. Các giao thức của người dùng cuối, chẳng hạn như NetBIOS, sẽ giới thiệu một rủi ro bảo mật lớn đối với hệ thống và giao thông trong DMZ.

• Cuộc tấn công zero-day
Các cuộc tấn công zero-day khai thác các lỗ hổng chưa được biết trước đó nên chúng đặc biệt nguy hiểm vì không thể thực hiện trước các biện pháp phòng ngừa. Một khoảng thời gian đáng kể có thể trôi qua kể từ khi một nhà nghiên cứu hoặc
kẻ tấn công phát hiện ra lỗ hổng và khi nhà cung cấp đưa ra bản vá khắc phục. Cho đến thời điểm đó, phần mềm
dễ bị tổn thương và thật không may là không có cách nào để chống lại những cuộc tấn công này. Để giảm thiểu thiệt hại, điều quan trọng là phải áp dụng các bản vá ngay khi chúng được phát hành.

• Đánh cắp xác thực

Để xác định người dùng, cá nhân hóa nội dung và đặt cấp độ truy cập, nhiều ứng dụng Web yêu cầu người dùng xác thực. Điều này có thể được thực hiện thông qua xác thực cơ bản (ID người dùng và mật khẩu), hoặc thông qua các phương thức xác thực mạnh hơn, chẳng hạn như yêu cầu chứng chỉ phía máy khách. Xác thực mạnh hơn có thể là cần thiết nếu yêu cầu chống chối bỏ.
Xác thực là thành phần chính của các dịch vụ xác thực, ủy quyền và kế toán (AAA) hầu hết các ứng dụng Web sử dụng. Như vậy, xác thực là tuyến phòng thủ đầu tiên để xác minh và theo dõi sử dụng hợp pháp một ứng dụng Web.
Một trong những vấn đề chính của việc xác thực là mọi ứng dụng Web đều thực hiện xác thực theo một cách khác. Việc thực thi chính sách xác thực nhất quán giữa nhiều ứng dụng và các ứng dụng khác nhau có thể tỏ ra đầy thách thức.
Việc chiếm quyền điều khiển xác thực có thể dẫn đến việc đánh cắp dịch vụ, chiếm quyền điều khiển phiên, mạo danh người dùng, tiết lộ thông tin nhạy cảm và leo thang đặc quyền. Kẻ tấn công có thể sử dụng các phương pháp xác thực yếu để giả định danh tính của người dùng khác và có thể xem và sửa đổi dữ liệu với tư cách là người dùng.

• Giả mạo nhật ký

Đầu tiên hãy kiểm tra xem trình duyệt Web có nhớ mật khẩu hay không. Các trình duyệt như Internet Explorer và Mozilla Firefox hỏi người dùng có nhớ mật khẩu hay không. Nếu người dùng quyết định làm điều này, mật khẩu đã lưu có thể bị đánh cắp. Một phương pháp khác để kiểm tra việc chiếm quyền xác thực là xem liệu người dùng có quên đăng xuất sau khi sử dụng hay không. Rõ ràng, nếu người dùng không đăng xuất, người tiếp theo sử dụng hệ thống có thể dễ dàng đặt ra với tư cách là người đó.

• Directory traversal

Các ứng dụng phức tạp tồn tại dưới nhiều thành phần và ứng dụng dữ liệu đặc biệt, thường được cấu hình trong nhiều thư mục. Một ứng dụng có khả năng duyệt qua nhiều thư mục này để định vị và thực thi những phần khác nhau của nó. Một cuộc tấn công truyền tải thư mục, còn được gọi là một cuộc tấn công tấn công duyệt web mạnh mẽ, xảy ra khi một Kẻ tấn công có thể duyệt các thư mục và tập tin bên ngoài quyền truy cập thông tin ứng dụng thông thường. Điều này làm rõ thư mục cấu trúc của một ứng dụng và thường là Web máy chủ và cơ sở điều hành.

• Chặn mật mã

Bằng cách sử dụng mật mã, một tin nhắn bí mật có thể được gửi một cách an toàn giữa hai bên. Sự phức tạp của ngày hôm nay. Các ứng dụng và cơ sở hạ tầng web thường liên quan đến nhiều điểm kiểm soát khác nhau nơi dữ liệu được mã hóa và được giải

mã. Ngoài ra, mọi hệ thống mã hoá hoặc giải mã thông điệp đều phải có bí mật cần thiết khóa và khả năng bảo vệ các khóa bí mật đó. Việc tiết lộ khóa riêng và chứng chỉ mang lại cho kẻ tấn công khả năng đọc và sửa đổi một giao tiếp riêng tư cho đến nay. Việc sử dụng mật mã và SSL phải được được xem xét cẩn thận vì lưu lượng được mã hóa đi qua tường lửa mạng và hệ thống IDS mà không được kiểm tra.
Nói cách khác, kẻ tấn công có một đường hầm được mã hóa an toàn để từ đó tấn công ứng dụng Web. Kẻ tấn công có thể chặn các tin nhắn được bảo mật bằng mật mã có thể đọc và sửa đổi các tin nhắn nhạy cảm, được mã hóa dữ liệu. Bằng cách sử dụng các khóa và chứng chỉ riêng đã bị bắt, kẻ tấn công trung gian có thể tàn phá hệ thống bảo mật, thường mà không làm cho các bên cuối cùng biết được điều gì đang xảy ra.

• URL interpretation

Tấn công giải thích URL là khi kẻ tấn công lợi dụng các phương pháp mã hóa văn bản khác nhau, lạm dụng việc giải thích URL. Bởi vì lưu lượng truy cập Web thường được hiểu là “thân thiện” nên nó không được lọc.
Đây là lưu lượng được sử dụng phổ biến nhất được phép thông qua tường lửa. Các URL được sử dụng cho kiểu tấn công này thường chứa các ký tự đặc biệt yêu cầu xử lý cú pháp đặc biệt để giải thích. Các ký tự đặc biệt thường được biểu thị bằng ký tự phần trăm, theo sau là hai chữ số biểu thị mã thập lục phân của số gốc
ký tự, tức là %<mã hex>. Bằng cách sử dụng các ký tự đặc biệt này, kẻ tấn công có thể đưa ra các lệnh độc hại hoặc nội dung, sau đó được thực thi bởi máy chủ Web. Một ví dụ về kiểu tấn công này là phản hồi HTTP chia nhỏ, trong đó kẻ tấn công có thể buộc hoặc chia một yêu cầu từ máy tính mục tiêu thành hai yêu cầu tới máy chủ.
Máy chủ web. Sau đó, kẻ tấn công tạo phản hồi gắn liền với một trong các yêu cầu máy chủ thực sự chứa dữ liệu giả mạo bởi kẻ tấn công. Dữ liệu giả mạo này sẽ được gửi trở lại mục tiêu, trông như thể nó đến trực tiếp từ máy chủ web.

• Tấn công mạo danh

Tấn công mạo danh là khi kẻ tấn công giả mạo các ứng dụng Web bằng cách giả vờ là người dùng hợp pháp. trong trường hợp này, kẻ tấn công vào phiên thông qua một cổng chung với tư cách là người dùng bình thường nên tường lửa không phát hiện được. Máy chủ có thể dễ bị tấn công bởi kiểu mã hóa quản lý phiên kém.
Quản lý phiên là một kỹ thuật sử dụng phiên để theo dõi thông tin. Các nhà phát triển web làm điều này để cung cấp ủy quyền minh bạch cho mọi yêu cầu HTTP mà không yêu cầu người dùng đăng nhập mỗi lần. Phiên tương tự như cookie ở chỗ chúng chỉ tồn tại cho đến khi bị hủy. Sau khi phiên bị hủy, trình duyệt ngừng mọi hoạt động theo dõi cho đến khi một phiên mới được bắt đầu trên trang Web. Ví dụ, giả sử ông A là

một người dùng hợp pháp và ông X là kẻ tấn công. Ông A duyệt đến một ứng dụng Web thương mại điện tử và cung cấp tên người dùng và mật khẩu, có được quyền truy cập hợp pháp vào thông tin, chẳng hạn như dữ liệu tài khoản ngân hàng của anh ta.

4. **Tổng quan về web logs**

Nguồn, tính chất và thời gian tấn công có thể được xác định bằng cách phân tích tệp nhật ký của hệ thống bị xâm nhập. Máy chủ Windows 2003 có các nhật ký sau:

• Nhật ký ứng dụng(application logs): lưu trữ các sự kiện liên quan đến ứng dụng đang chạy trên máy chủ

• Nhật ký bảo mật(security logs): lưu trữ các sự kiện liên quan đến kiểm toán

• Nhật ký hệ thống(system log): lưu trữ các sự kiện liên quan đến các thành phần và dịch vụ của Windows

• Nhật ký Dịch vụ Thư mục, lưu trữ thông tin lỗi và chẩn đoán Active Directory

• Nhật ký Dịch vụ sao chép (Directory Service log): lưu trữ các sự kiện sao chép tệp Active Directory

• Nhật ký dành riêng cho dịch vụ, lưu trữ các sự kiện liên quan đến dịch vụ hoặc ứng dụng cụ thể

Tệp nhật ký có mã trạng thái HTTP dành riêng cho loại sự cố. Mã trạng thái được chỉ định trong HTTP và phổ biến cho tất cả các máy chủ Web. Mã trạng thái là số có ba chữ số trong đó chữ số đầu tiên xác định
lớp phản ứng. Mã trạng thái được phân thành năm loại.

Không cần thiết phải hiểu định nghĩa của các mã trạng thái HTTP cụ thể mà điều quan trọng là phải hiểu hiểu lớp của mã trạng thái. Bất kỳ mã trạng thái nào của một lớp phải được xử lý giống như bất kỳ mã trạng thái nào những người khác thuộc lớp đó.

Status Code	Description
1XX	Continue or request received
2XX	Success
3XX	Redirection
4XX	Client error
5XX	Server error

Hình 4: Status Code của Log Windows

5. Quy trình điều tra tấn công web

Để điều tra các cuộc tấn công Web, điều tra viên nên làm theo các bước sau:

- Phân tích máy chủ Web, máy chủ FTP và nhật ký hệ thống cục bộ để xác nhận một cuộc tấn công Web.
- Kiểm tra thông tin tệp nhật ký liên quan đến dấu thời gian, địa chỉ IP, mã trạng thái HTTP và tài nguyên được yêu cầu.

- Xác định bản chất của cuộc tấn công. Điều cần thiết là phải hiểu bản chất của cuộc tấn công; nếu không, nó sẽ khó có thể ngăn chặn nó ở giai đoạn đầu. Nếu không dừng lại sớm, nó có thể vượt khỏi tầm tay.
  - Kiểm tra xem có ai đó đang cố tắt mạng hoặc đang cố xâm nhập vào hệ thống hay không.
  - Trực quan hóa nguồn.
  - Sử dụng tường lửa và nhật ký IDS để xác định nguồn tấn công. IDS và tường lửa giám sát lưu lượng mạng và ghi lại từng mục. Những điều này giúp xác định xem nguồn tấn công có phải là máy chủ bị xâm nhập hay không trên mạng hoặc bên thứ ba.
- Chặn cuộc tấn công. Sau khi xác định được cách kẻ tấn công đã xâm nhập vào hệ thống, cổng hoặc lỗ hổng đó sẽ bị chặn để ngăn chặn sự xâm nhập tiếp theo.
  - Sau khi xác định được hệ thống bị xâm nhập, hãy ngắt kết nối chúng khỏi mạng cho đến khi chúng có thể được khử trùng. Nếu cuộc tấn công đến từ nguồn bên ngoài, hãy chặn ngay địa chỉ IP đó.
  - Bắt đầu điều tra từ địa chỉ IP.

6. Đánh giá đề tài

- Đánh giá

  Độ khó chuyên đề: Chuyên đề nằm ở mức độ trung bình, có thể tài liệu tham khảo.

  - Mức độ đọc hiểu tài liệu liên quan đến chuyên đề : Đọc hiểu được tài liệu liên quan, khó trong việc chọn lọc tài liệu đúng.
  - Thời gian thực hiện : Phù hợp với chế độ làm việc nhóm, chuyên đề không tốn Qúa nhiều phí và thời gian.
  - Mức độ tối ưu khi xây dựng chương trình : đánh giá ở phần mục đánh giá kết quả thực hiện chương trình ở cuối chuyên đề.

- Định hướng đề tài

  - Chuyên đề lựa chọn đòi hỏi tính tối ưu khi xây dựng môi trường thực nghiệm, khả năng khai thác và sử dụng nguồn tài liệu có sẵn.
  - Phải hiểu rõ được các cơ chế hoạt động, tính thiết yếu khi đi vào tấn công thực nghiệm.

- Xác định tầm quan trọng của chuyên đề, từ đó nâng cao cải thiện cho chương trình.

- Tiến hành thực hiện điều tra đối với các lỗ hổng khác

- Khai thác tìm kiếm lỗ hổng trong cơ chế hoạt động tấn công, từ đó phân tích và đúc kết hướng đi cho việc thiết lập phòng chống.

- Bám sát vào bố cục ý tưởng của tài liệu tham khảo, thực hiện tấn công thực nghiệm theo từng ý tưởng một.

Chương II: Triển khai thực nghiệm và đánh giá

2.1 Xây dựng hệ thống để chuẩn bị tấn công

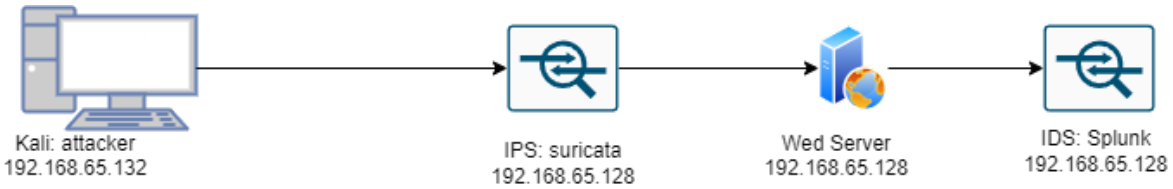
- Trong kịch bản tấn công này ta sẽ tiến hành điều tra một trang web bị tấn công SQL Injection. Chính vì thế nên ta sẽ xây dựng một hệ thống gồm những thứ sau:

+ Một máy để xây dựng trang web có lỗ hổng và dựng suricata làm IPS

+ Một máy xây dựng hệ thống thu thập log tập trung splunk

- Kịch bản sẽ tiến hành như sau:

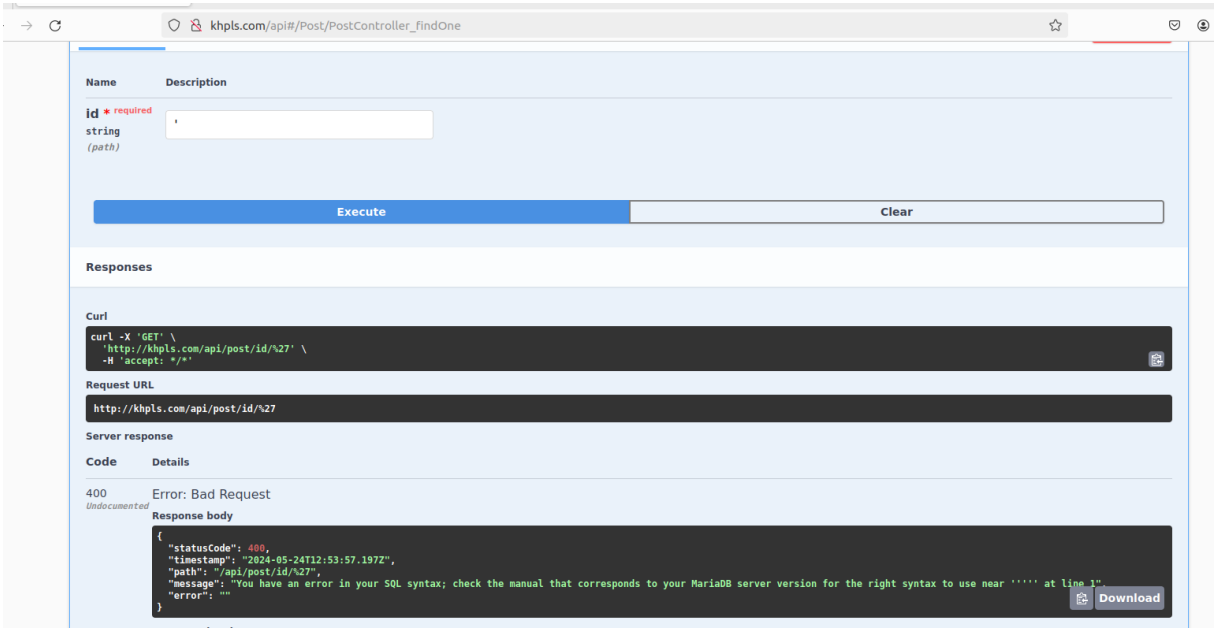
Khi attacker tiến hành tấn công SQL injection thông qua trang web thì sẽ đi qua surikata và vào trong trang web. Khi này suricata phát hiện và sẽ chặn truy cập của máy có ip đó. Khi đó, log của web server sẽ được chuyển về máy có chứa splunk. Máy splunk sẽ kích hoạt alert khi phát hiện các cú pháp của tấn công sql injection và sẽ gửi mail về cảnh báo là có tấn công. Chúng ta sẽ dựa vào nguồn log đó và điều tra cuộc tấn công này.



Hình 5: Mô hình của bài Lab

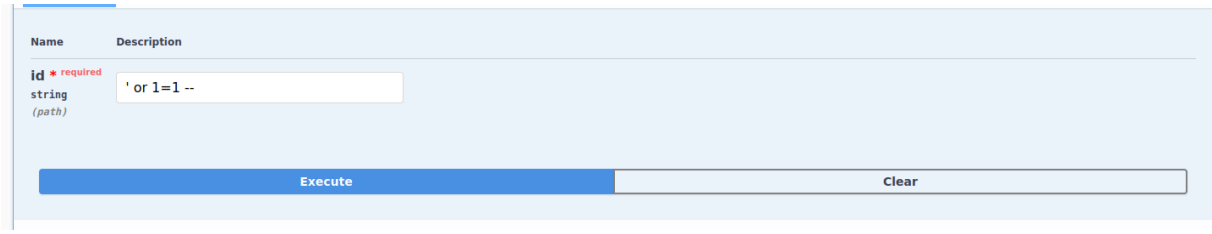
2.2 Bắt đầu thực nghiệm

Ở đây chúng ta sẽ có một trang API của trang web chúng ta sẽ dùng để thực nghiệm. Ở đây cho phép chúng ta điền các tham số SQl để nó truy vấn và gửi kết quả xuống bên dưới. Như hình chúng ta có thể thấy đây là phần truy vấn id của các bài post. Chúng ta chỉ cần điền id vào và trang web sẽ trả lại cho chúng ta nội dung trong Database tương ứng với trường chúng ta nhập vào.



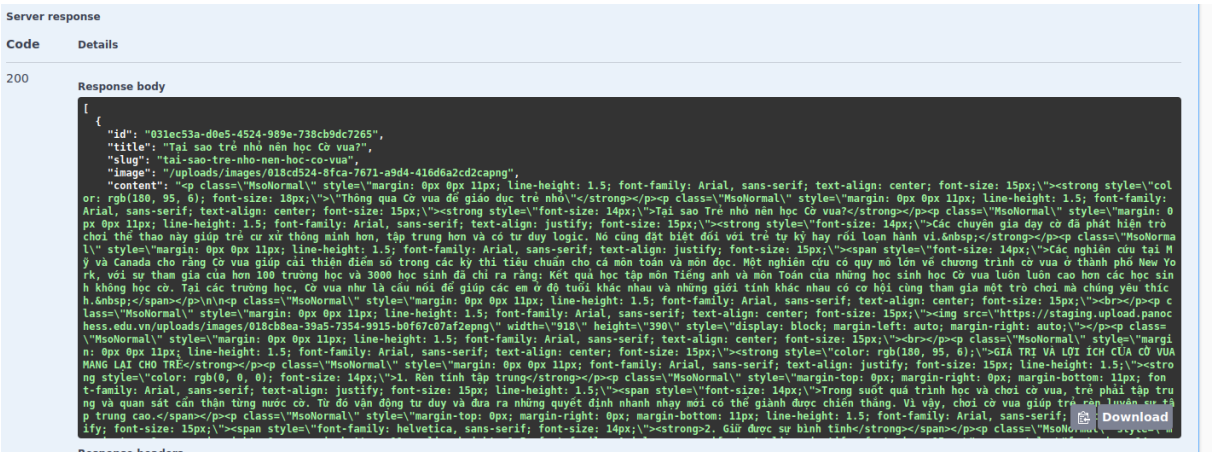
Hình 6: Xác nhận trang Web bị nhiễm SQL Injection

Tuy nhiên, khi đưa kí tự ‘ vào trong trường này thì trang web bị lỗi. Điều này chứng tỏ rằng trang web bị nhiễm SQL injection



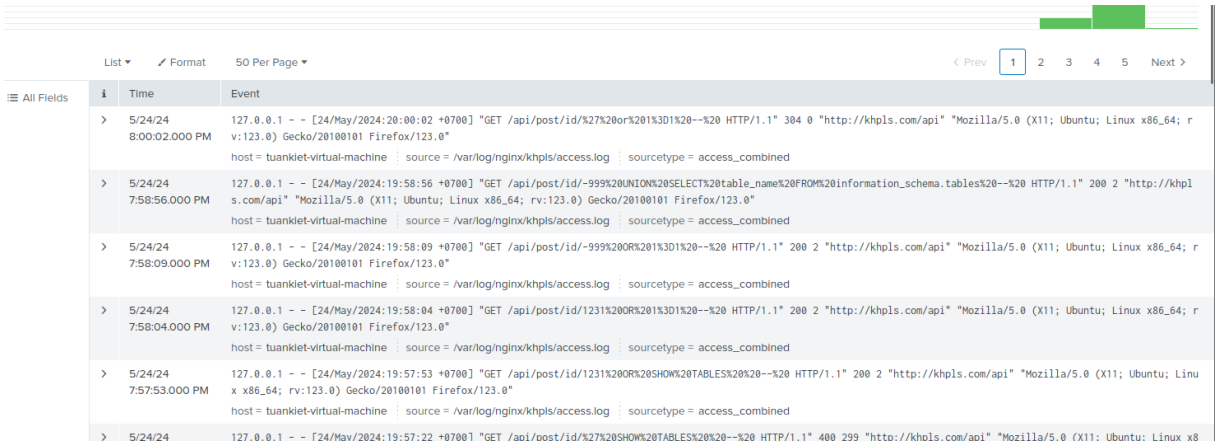
Hình 7: Thực hiện SQL Injection

Ta đưa câu lệnh muốn SQL thực hiện vào trong trường id và ta sẽ nhận được những thông tin sau

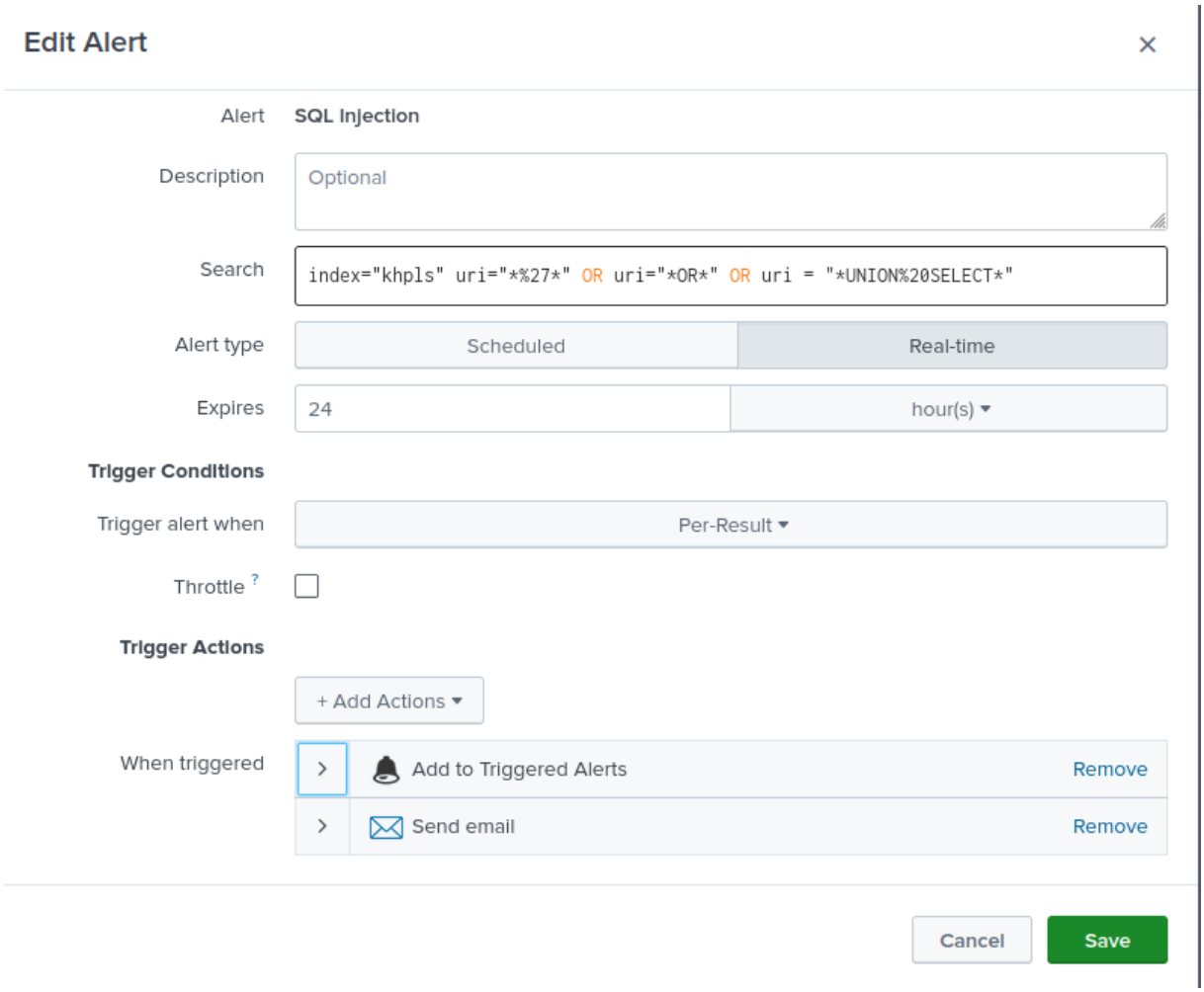


Hình 8: Kết quả khi thực hiện SQL Injection

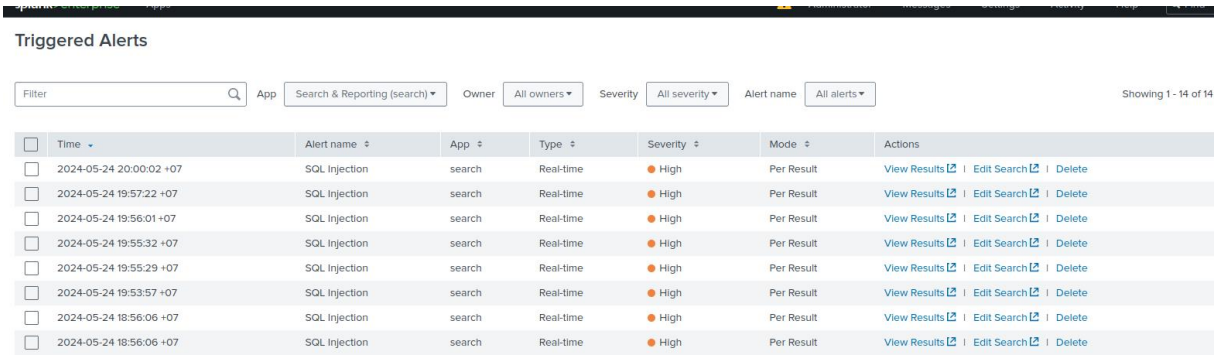
Ta lấy được tất cả các thông tin trong table đó. Ta đã tấn công SQL Injection thành công. Trong bài lab này, ta có cài đặt thêm splunk. Nó sẽ làm một trung tâm thu thập log. Nói cách khác, mọi câu truy vấn mà chúng ta truy vấn ở đây đều được đưa về splunk để hiển thị và lưu trữ. Chúng ta làm vậy để hiển thị log tập tủng cũng như là tránh trường hợp log bị xoá bởi hacker.



Hình 9: Log thu thập được bởi splunk

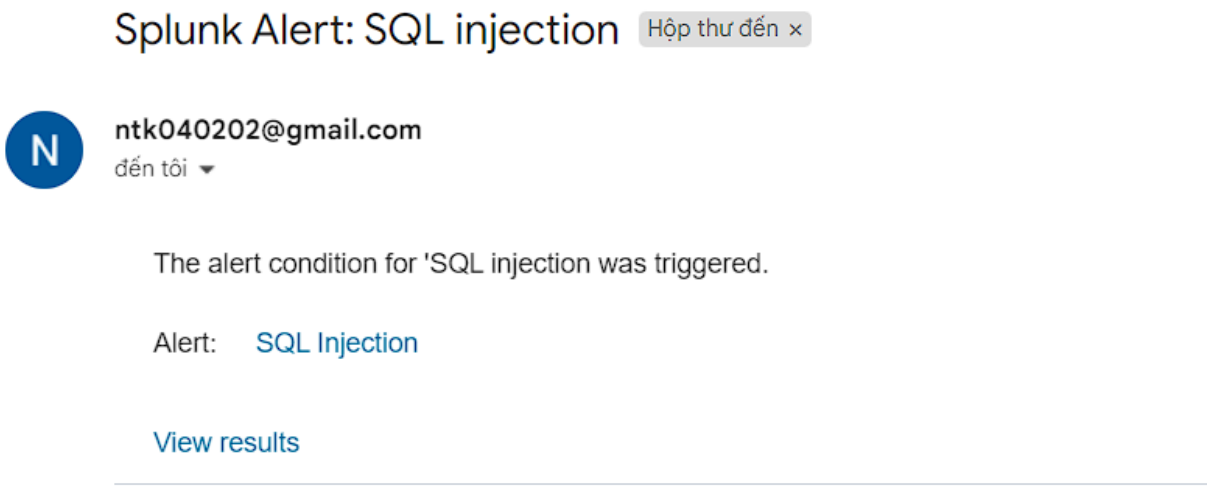


Hình 10: Cài đặt alert SQL Injection



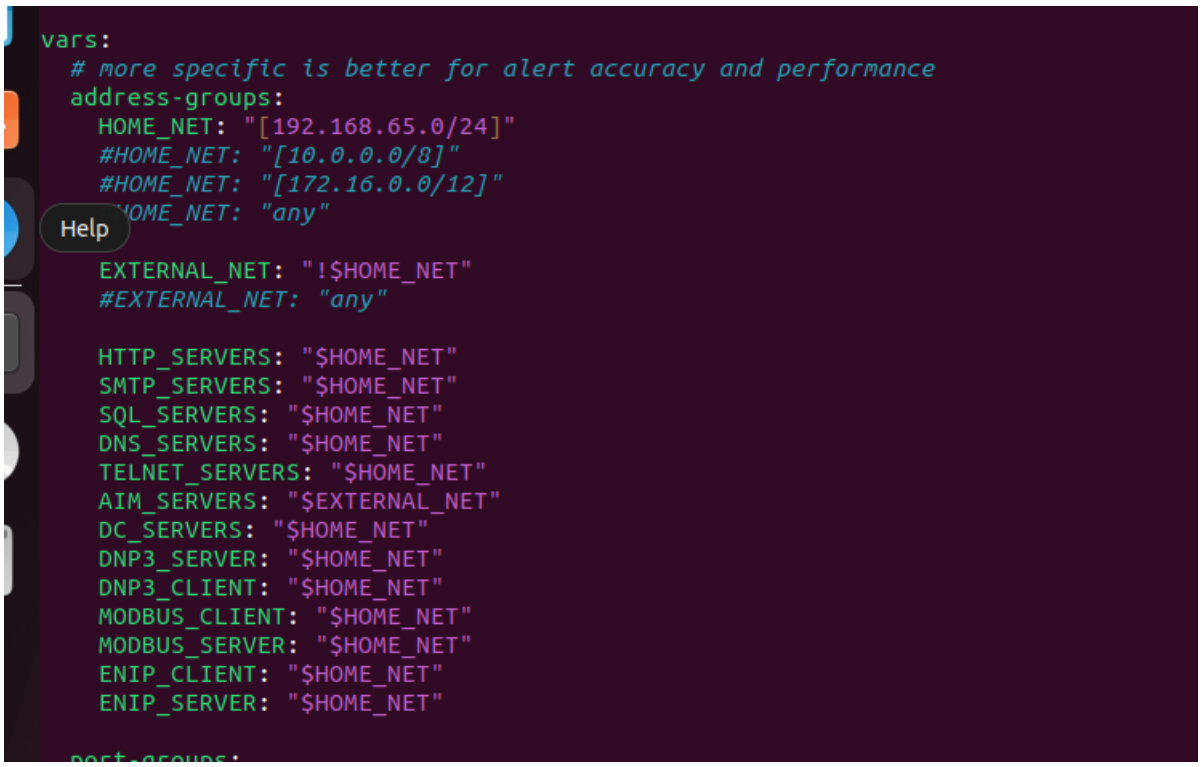
Hình 11: Alert của Splunk

Hiển nhiên trên splunk cũng có cài đặt alert để cảnh báo khi nó phát hiện bị tấn công SQL Injection. Kết quả của alert là sẽ được hiển thị lên màn hình alert và gửi gmail về cho người quản trị



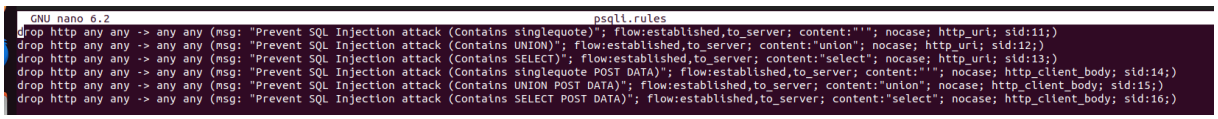
Hình 12: Gmail được Splunk gửi về

Trong suricata được chúng ta thiết lập để phát hiện tấn công SQl Injection thì sẽ tự động chặn IP của kẻ tấn công. Để làm được điều đó thì ta phải thiết lập rule cho Suricata và tiến hành thực thi nó



Hình 13: Cấu hình HOME_NET suricata

Hình 14: Thêm đường dẫn tới rule trong suricata



Hình 15: Rule cảu suricata

Trong đoạn rule trên, chúng ta sẽ drop các gói tin nếu phát hiện chúng có các kí tự sau trong câu truy vấn:

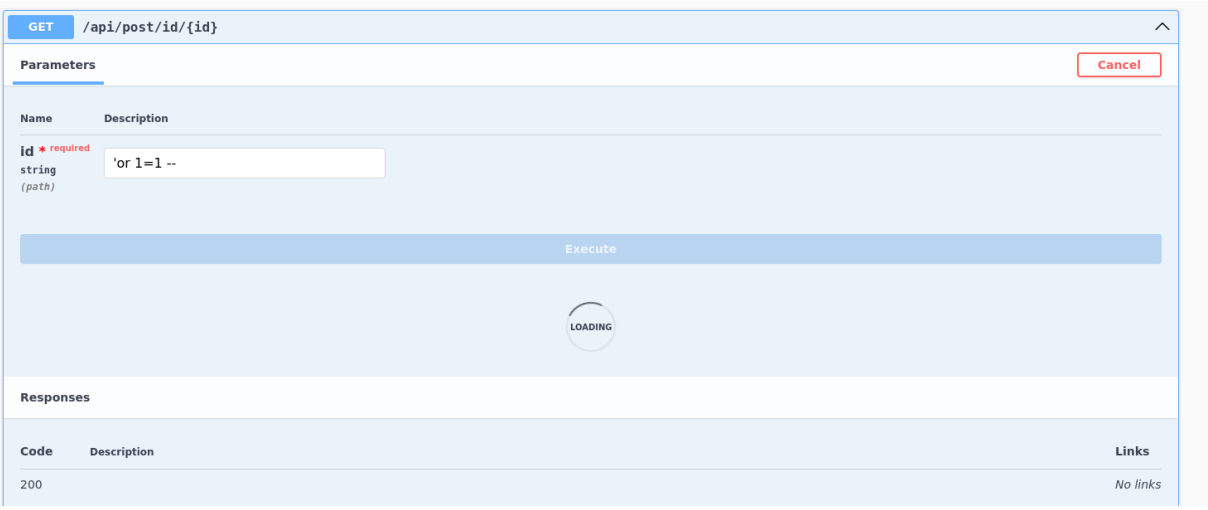
- nháy đơn

- nháy kép

- UNION

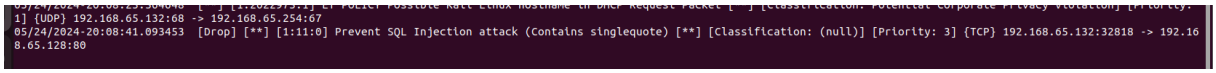
Hình 16: IPS Mode của IPS

Sau khi ta tiến hành kích hoạt chế độ IPS thì ta sẽ tiến hành tấn công lại SQL Injection vào trang web để xem thử có còn được không



Hình 17: Tấn công SQL lần 2

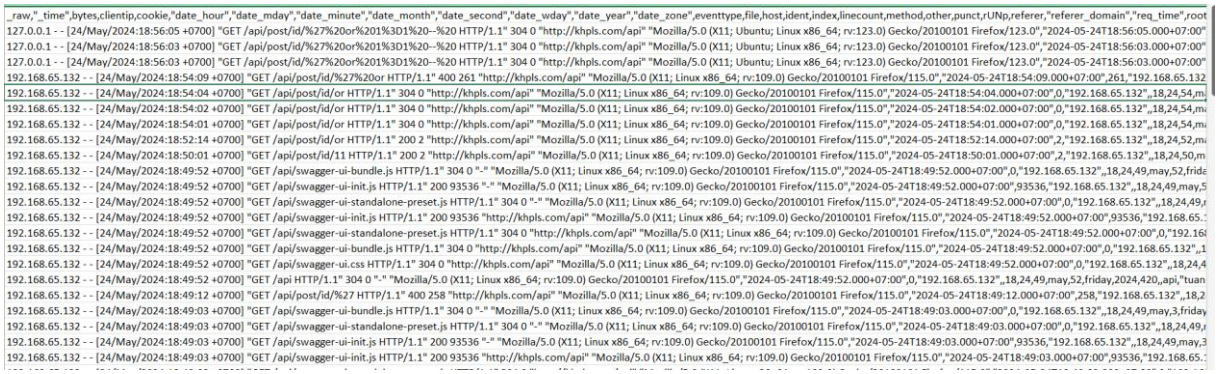
Lần này thì chúng ta tấn công không được nữa do gói tin đã bị chặn bởi suricata do trong câu truy vấn của chúng ta có chứa các kí tự nằm trong rule của suricata.



Hình 18: Thông báo đã chặn gói tin

2.3 Tiến hành điều tra

Chúng ta đã chặn được cuộc tấn công. Bây giờ ta sẽ tiến hành dựa vào log để điều tra xem ai là kẻ tấn công và kẻ tấn công đó đã lấy đi thông tin gì của chúng ta. Từ Splunk ta sẽ xuất file chứa đựng thông tin các log mà ta đã thu thập được và xem xét các log được splunk ghi lại khi cuộc tấn công xảy ra.



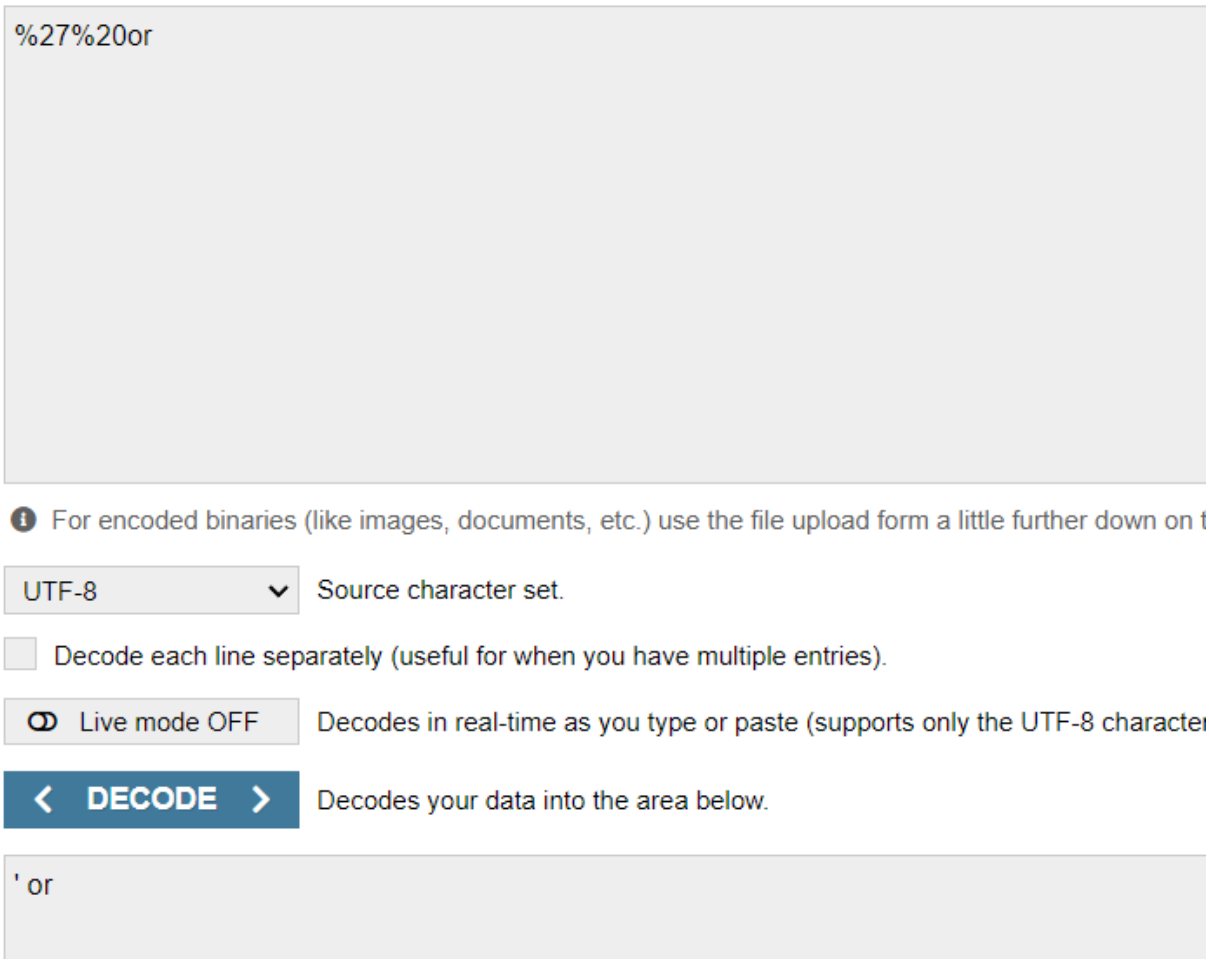
Hình 19: Log lấy từ splunk

Bây giờ ta sẽ tiến hành dựa vào log này và lấy những thông tin cần thiết

Hình 20: Log tấn công ngày 24

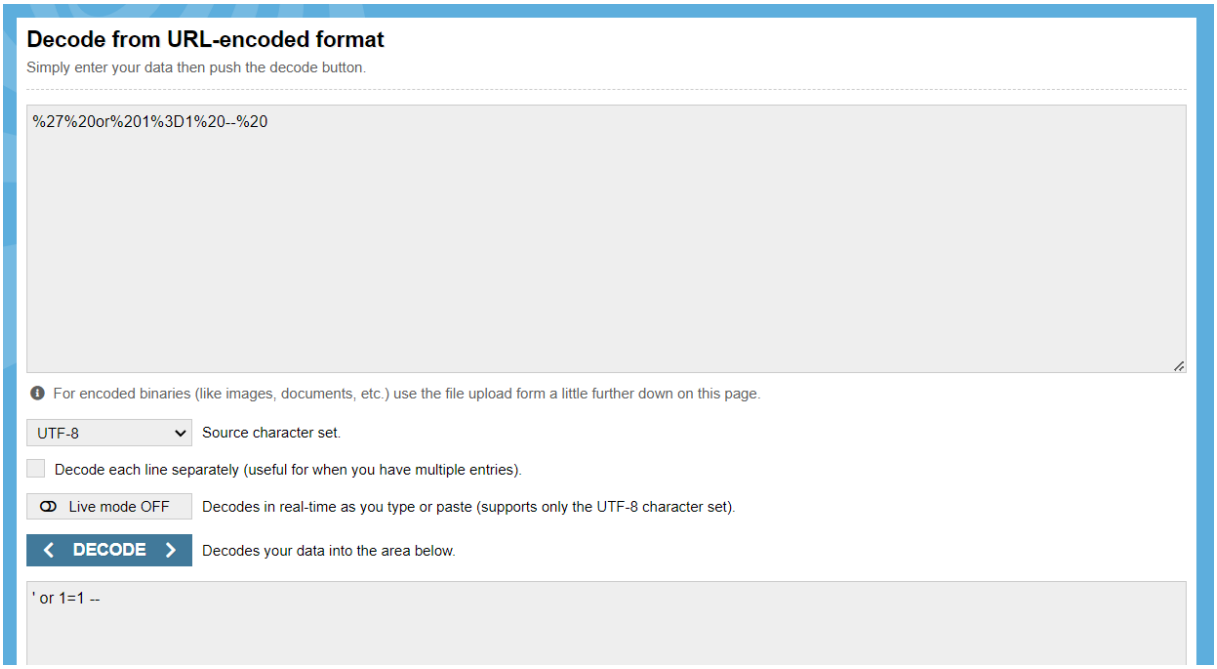
Trong ngày 24 này ta có một vụ tấn công xảy ra thì ta có được các thông tin hữu ích sau:

- Ip máy tấn công: 192.168.65.132
- Truy cập từ lúc 18 giờ 49 và tấn công vào lúc 18 giờ 54 thì bị chặn mất
- Hacker dùng trình duyệt Mozilla phiên bản 115 để tấn công trang web của chúng ta
- Câu lệnh được hacker truy vấn là ‘or. Hacker chỉ vừa mới kiểm thửu xem có bị SQL Injection không mà đã bị chặn mất
- Ngoài ra chúng ta còn biết thêm đó là hacker tấn công vào mục post id. Mục vào nhằm mục đích là truy vấn các bài viết nhưng không được kiểm tra kĩ càng nên đã bị dính lỗi SQL Injection



Hình 21: Decode URL

Hình 22: Thử SQL Injection



Hình 23: Decode URL ngày 24

Tuy nhiên khi kéo xuống dưới thì ta đã thấy ngày 23 thì kẻ tấn công cũng đã tấn công chúng ta một lần. Lúc này kẻ tấn công sử dụng sqlmap để dò quét.

Hình 24:Quét bằng SQL Injection

Với sự dò quét của SQl map có thể kẻ tấn công đã có trong tay toàn bộ Database của chúng ta.

Bây giờ ta đã xác nhận được thiệt hại cũng như con đường mà hacker đã xâm nhập. ta tiến hành sửa lỗi và cho trang web hoạt động lại bình thường.

IP tấn công chúng ta lần này là IP nội bộ. Việc bị một máy IP nội bộ tấn công chứng tỏ rằng hacker đã xâm nhập vào trong hệ thống nội bộ và kiểm soát máy có IP 192.168.65.132 và tấn công web server của chúng ta. Bây giờ muốn điều tra sâu hơn về IP đó thì ta phải tiến hành tiếp cận và điều tra trên máy Windows.

CHƯƠNG III: KẾT QUẢ VÀ PHÂN TÍCH

3.1 Kết quả

Sau quá trình điều tra kỹ lưỡng, chúng tôi đã thu thập và phân tích các thông tin chi tiết liên quan đến cuộc tấn công web . Kết quả của cuộc điều tra đã tiết lộ một loạt các hoạt động không đáng tin cậy trên hệ thống mạng, từ việc chạy các yêu cầu HTTP đến việc thử nghiệm các lỗ hổng bảo mật.

Trong số các phát hiện đáng chú ý, một số điểm cần được nêu ra bao gồm:

Phát hiện của các yêu cầu HTTP không hợp lệ: Chúng tôi đã ghi nhận một lượng lớn các yêu cầu HTTP đến các tài nguyên không tồn tại trên hệ thống. Điều này thường được coi là một biểu hiện của các nỗ lực tấn công kiểu lùi.

Kiểm tra thử các lỗ hổng bảo mật: Các log cho thấy một số yêu cầu HTTP đặc biệt được thiết kế để thử nghiệm các lỗ hổng bảo mật cụ thể trên hệ thống. Các yêu cầu này thường liên quan đến việc kiểm tra điều kiện đầu vào và thực hiện các kỹ thuật tấn công SQL Injection

Các giao tiếp không an toàn giữa máy chủ và client: Chúng tôi cũng phát hiện một số yêu cầu HTTP được gửi qua kênh không an toàn, không mã hóa, có thể làm lộ thông tin quan trọng như thông tin đăng nhập hoặc dữ liệu nhạy cảm khác.

Tóm lại, kết quả của điều tra tấn công web đã cung cấp cái nhìn chi tiết về các hoạt động không đáng tin cậy trên hệ thống mạng, cũng như những lỗ hổng bảo mật tiềm ẩn mà cần phải khắc phục ngay lập tức để bảo vệ hệ thống khỏi các mối đe dọa tiềm ẩn.

3.2 Phương hướng phát triển trong tương lai

- Hiện tại hệ thống IPS chỉ có thể chặn những cuộc tấn công về SQL Injection và chưa thể ngăn chặn các cuộc tấn công khác do chưa cập nhật rule cho hệ thống. Chúng em sẽ cập nhật rule mới cho hệ thống trong tương lai

- Splunk cũng mới chỉ alert khi phát hiện SQL Injection mà chưa phát hiện thêm các mỗi đe doạ khác. Trong thười gian tới sẽ cập nhật them cho các loại tấn công khác để bảo vệ web một cách tốt hơn.

Tài liệu tham khảo

https://fptshop.com.vn/tin-tuc/danh-gia/website-la-gi-169249

https://bizflycloud.vn/tin-tuc/web-application-la-gi-co-gi-khac-voi-website-20180619105652369.htm

https://www.linkedin.com/pulse/t%E1%BA%A5n-c%C3%B4ng-%E1%BB%A9ng-d%E1%BB%A5ng-web-l%C3%A0-g%C3%AC-v%C3%A0-c%C3%A1ch-b%E1%BA%A3o-v%E1%BB%87-what-application-cao

Phân công công việc

STT	MSSV	Họ và tên	Công việc
1	N20DCAT040	Dương Minh Phong	+ Xây dựng hệ thống IPS để phát hiện web bị tấn công. + Sử dụng nguồn log để điều tra cuộc tấn công
2	N20DCAT048	Phan Tiến Sĩ	Sử dụng nguồn log có sẵn thực hiện điều tra Build splunk để sự trữ tập trung log
3	N20DCAT028	Nguyễn Tuấn Kiệt	+ Sử dụng các công cụ để tấn công web nhằm tạo ra một vụ tấn công + Sử dụng nguồn log để điều tra vụ tấn công
4	N20DCAT057	Phan Công Thăng	Dựng một webapp có chứa lỗ hổng