BỘ CÔNG THƯƠNG

TRƯỜNG ĐẠI HỌC CÔNG THƯƠNG TP. HCM

KHOA CÔNG NGHỆ THÔNG TIN

ĐỒ ÁN CHUYÊN NGÀNH

TRIỂN KHAI HỆ THỐNG SIEM
ỨNG DỤNG TRONG PHÁT HIỆN TẤN CÔNG

GVHD: TS Vũ Đức Thịnh

SVTH:

Nguyễn Thị Thu Hoa 2033210282 –12DHBM1

Trần Minh Nhựt 2033210621 –12DHBM1

Phạm Hoàng Bảo 2033216354 –12DHBM2

Thành phố Hồ Chí Minh, tháng 12 năm 2024

BỘ CÔNG THƯƠNG

TRƯỜNG ĐẠI HỌC CÔNG THƯƠNG TP. HCM

KHOA CÔNG NGHỆ THÔNG TIN

ĐỒ ÁN CHUYÊN NGÀNH

TRIỂN KHAI HỆ THỐNG SIEM
ỨNG DỤNG TRONG PHÁT HIỆN TẤN CÔNG

GVHD: TS Vũ Đức Thịnh

SVTH:

Nguyễn Thị Thu Hoa 2033210282 –12DHBM1

Trần Minh Nhựt 2033210621 –12DHBM1

Phạm Hoàng Bảo 2033216354 –12DHBM2

Thành phố Hồ Chí Minh, tháng 12 năm 2024

LỜI CAM ĐOAN

Chúng tôi xin cam đoan đây là công trình nghiên cứu của riêng chúng tôi. Các số liệu, kết quả nêu trong Đồ án là trung thực và chưa từng được ai công bố trong bất kỳ công trình nào khác.

Tôi xin cam đoan rằng mọi sự giúp đỡ cho việc thực hiện Đồ án này đã được cảm ơn và các thông tin trích dẫn trong Đồ án đã được chỉ rõ nguồn gốc.

Sinh viên thực hiện Đồ án

(Ký và ghi rõ họ tên)

LỜI CÁM ƠN

Trước tiên, chúng em muốn gửi lời cảm ơn và biết ơn chân thành đến thầy Vũ Đức Thịnh, người đã hướng dẫn chúng em trong quá trình thực hiện đề tài. Thầy đã tận tình chỉ bảo và hỗ trợ nhóm suốt thời gian thực hiện, cũng như đóng góp ý tưởng và kiểm tra tính phù hợp của đề tài.

Chúng em cũng muốn gửi lời cảm ơn đến toàn thể các thầy cô giáo của Trường ĐH Công Thương TP.HCM, vì đã truyền đạt kiến thức và tạo điều kiện thuận lợi cho chúng em trong quá trình học tập và phát triển tại trường.

Mặc dù chúng em đã cố gắng hoàn thành đề tài trong phạm vi và khả năng của mình, tuy nhiên không thể tránh khỏi những thiếu sót. Chúng em rất mong nhận được sự cảm thông và sự hướng dẫn tận tâm từ quý thầy cô.

Xin chân thành cảm ơn!

(Họ và tên của Tác giả Đồ án)

TÓM TẮT

Triển khai hệ thống SIEM (Security Information and Event Management) là một biện pháp quan trọng trong việc tăng cường an ninh mạng và phát hiện các cuộc tấn công tiềm ẩn. SIEM thu thập, phân tích và lưu trữ log từ nhiều nguồn khác nhau như tường lửa, máy chủ, và các ứng dụng, giúp nhận diện các hành vi bất thường và cảnh báo sớm về các mối đe dọa như tấn công DoS, SQL Injection, Brute Force. Hệ thống cung cấp khả năng cảnh báo kịp thời, giúp đội ngũ an ninh phản ứng nhanh chóng, hiệu quả với sự cố và hỗ trợ tổ chức tuân thủ các quy định pháp lý. Mặc dù triển khai SIEM có thể phức tạp và tốn kém, nhưng lợi ích về phát hiện sớm, giảm thiểu rủi ro và cải thiện khả năng ứng phó sự cố giúp bảo vệ tối đa an toàn hệ thống thông tin của tổ chức.

MỤC LỤC

Danh mục hình ảnh

Danh mục bảng biểu

DANH MỤC CÁC TỪ VIẾT TẮT

STT	Kỹ tự chữ viết tắt	Chữ viết đầy đủ
1	SIEM	Security information and event management
2	IDS	Intrusion Detection System
3	IPS	Intrusion Prevention System
4	ML	Machine Learning
5	MLTK	Machine Learning Toolkit
6	CNTT	Công nghệ thông tin
7	OWASP	Open Web Application Security Project

ĐÁNH GIÁ VÀ NHẬN XÉT CỦA GIẢNG VIÊN HƯỚNG DẪN

.

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

Giảng viên hướng dẫn (Kí và ghi rõ họ tên)

ĐÁNH GIÁ VÀ NHẬN XÉT CỦA GIẢNG VIÊN PHẢN BIỆN

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

.....

Giảng viên phản biện (Kí và ghi rõ họ tên)

MỞ ĐẦU

1. Lý do chọn đề tài

Các doanh nghiệp tại Việt Nam đang đối mặt với nhiều thách thức an ninh mạng do tốc độ số hóa nhanh chóng và thiếu chú trọng bảo mật thông tin, dẫn đến nguy cơ mất dữ liệu và lộ thông tin nhạy cảm. Việc sử dụng IoT, điện toán đám mây và các nền tảng công nghệ mới tạo ra nhiều lỗ hổng, đồng thời đòi hỏi doanh nghiệp phải nâng cao quản lý an ninh mạng, đặc biệt trong bối cảnh các quy định như GDPR và Nghị định 85/2016/NĐ-CP ngày càng chặt chẽ.

Hệ thống SIEM (Security Information and Event Management) là một giải pháp tiên tiến, giúp doanh nghiệp thu thập, phân tích sự kiện và phản hồi kịp thời các sự cố an ninh, đồng thời hỗ trợ tuân thủ pháp luật thông qua khả năng lưu trữ và báo cáo dữ liệu. Nhận thấy tầm quan trọng này, chúng em chọn đề tài "Triển khai hệ thống SIEM ứng dụng trong phát hiện tấn công mạng" nhằm góp phần nâng cao bảo mật thông tin cho doanh nghiệp, đồng thời áp dụng kiến thức vào thực tiễn.

2. Mục tiêu đề tài

Đề tài "Triển khai hệ thống SIEM ứng dụng trong phát hiện tấn công mạng" nhằm nghiên cứu khả năng, lợi ích và thách thức của hệ thống SIEM, đồng thời đánh giá sự hiệu quả của nó trong việc nâng cao khả năng phát hiện các mối đe dọa an ninh mạng và ứng phó với các sự cố. Các mục tiêu cụ thể bao gồm:

a. **Đánh giá tính năng thu thập và phân tích log của SIEM**
   - Tìm hiểu cơ chế thu thập log từ các hệ thống, thiết bị khác nhau (firewall, máy chủ, hệ điều hành, ứng dụng).
   - Đánh giá khả năng xử lý và phân tích log theo thời gian thực, từ đó hỗ trợ quá trình phát hiện và phản hồi các sự cố an ninh mạng nhanh chóng, hiệu quả.
   - Xác định các phương pháp tổ chức và quản lý log để đảm bảo tính toàn vẹn và bảo mật của dữ liệu.

b. **Nghiên cứu các cuộc tấn công mạng thực tế và vai trò của SIEM trong phát hiện và ngăn chặn tấn công**

- Phân tích các trường hợp tấn công mạng đã xảy ra (tấn công DDoS, xâm nhập trái phép, tấn công qua lỗ hổng bảo mật, malware).
- Tìm hiểu cách hệ thống SIEM phát hiện các dấu hiệu của những cuộc tấn công này thông qua phân tích hành vi bất thường và các mẫu sự kiện.
- Đánh giá khả năng của SIEM trong việc tự động hóa quá trình cảnh báo và phản hồi sự cố, giúp giảm thiểu thời gian và nguồn lực cần thiết để xử lý tấn công.

c. Đánh giá hiệu quả của chiến lược triển khai SIEM

- Xem xét các mô hình triển khai SIEM phù hợp với quy mô và nhu cầu của doanh nghiệp (on-premises, cloud-based hoặc hybrid).
- Đánh giá hiệu quả của SIEM trong việc đảm bảo tính liên tục của hoạt động kinh doanh và giảm thiểu rủi ro mất dữ liệu, gián đoạn hệ thống.
- Tìm hiểu các thách thức trong việc triển khai và vận hành hệ thống SIEM, bao gồm chi phí, yêu cầu về nguồn lực và nhân sự, và các rào cản kỹ thuật như việc tích hợp với các hệ thống hiện có.

d. Nghiên cứu các tiêu chuẩn và quy định liên quan đến triển khai SIEM

- Phân tích vai trò của SIEM trong việc giúp doanh nghiệp tuân thủ các tiêu chuẩn an ninh thông tin như ISO 27001, PCI-DSS, GDPR.
- Đánh giá hiệu quả của SIEM trong việc cung cấp báo cáo kiểm toán và hỗ trợ điều tra sau sự cố an ninh.

e. Đề xuất giải pháp cải thiện và tối ưu hóa hệ thống SIEM

- Đề xuất các biện pháp nhằm tối ưu hóa quá trình triển khai và vận hành hệ thống SIEM để đạt được hiệu suất cao nhất, giảm thiểu cảnh báo giả và tăng cường khả năng phát hiện các mối đe dọa mới.
- Đánh giá xu hướng phát triển của SIEM và các công nghệ bổ sung như Machine Learning trong việc nâng cao khả năng phát hiện tấn công và phản ứng tự động.

3. Đối tượng và phạm vi nghiên cứu

a. Đối tượng nghiên cứu

Đề tài tập trung nghiên cứu hệ thống SIEM (Security Information and Event Management) và các mối đe dọa từ các phương thức tấn công mạng. Đối tượng chính bao gồm:

- Các thành phần, tính năng và chức năng của hệ thống SIEM.
- Các phương thức tấn công mạng phổ biến và có thể triển khai trong môi trường học tập, bao gồm:
  - Tấn công từ chối dịch vụ (DoS): Làm quá tải tài nguyên hệ thống và khiến dịch vụ không thể truy cập.
  - Tấn công SQL Injection: Chèn mã SQL độc hại vào các truy vấn của cơ sở dữ liệu, dẫn đến việc truy cập hoặc thay đổi dữ liệu trái phép.
  - Tấn công brute-force: Thử nhiều tổ hợp mật khẩu khác nhau để truy cập vào hệ thống.

b. Phạm vi nghiên cứu

Đề tài sẽ bao quát các khía cạnh công nghệ và thực tế trong phạm vi triển khai và vận hành hệ thống SIEM, với trọng tâm nghiên cứu và phân tích các phương thức tấn công mạng phổ biến. Cụ thể bao gồm:

❖ **Phạm vi công nghệ SIEM**

- Nghiên cứu cơ chế thu thập, phân tích và lưu trữ log từ các thiết bị mạng, máy chủ, và ứng dụng, cũng như tích hợp SIEM với các công cụ an ninh khác.
- Khả năng giám sát, phát hiện và phản ứng nhanh chóng với các sự kiện an ninh mạng dựa trên phân tích log.
- Đánh giá tính hiệu quả của SIEM trong việc bảo vệ hệ thống mạng khỏi các cuộc tấn công phổ biến như DoS và SQL Injection trong các môi trường học tập hoặc phòng thí nghiệm.

❖ **Phạm vi các mối đe dọa và tấn công mạng**

Phân tích và triển khai các phương thức tấn công như:

- DoS: Tạo quá tải cho hệ thống để làm gián đoạn dịch vụ.
- Bruce Force: Password, địa chỉ các tệp thư mục.

Tìm hiểu cách thức hệ thống SIEM có thể phát hiện và ngăn chặn các cuộc tấn công này thông qua phân tích hành vi bất thường và log hệ thống.

❖ **Giới hạn nghiên cứu**

Nghiên cứu giới hạn trong phạm vi các phương thức tấn công mạng phổ biến và có thể thực hiện trong môi trường học tập, không bao gồm các công nghệ nâng

cao như Machine Learning. Việc triển khai SIEM sẽ được xem xét dựa trên môi trường thí nghiệm và các tổ chức học thuật.

4. Phương pháp nghiên cứu

Đề tài sử dụng các phương pháp nghiên cứu chính sau:

a. Phân tích các giải pháp SIEM:

Nghiên cứu các hệ thống SIEM phổ biến trên thị trường như Splunk, IBM QRadar, Elastic SIEM, ArcSight dựa trên các tiêu chí:

- Tính năng: Bao gồm khả năng thu thập và phân tích log, phát hiện tấn công theo thời gian thực, tính năng tự động hóa trong phản hồi sự cố, và các công cụ trực quan hóa dữ liệu.
- Khả năng tích hợp: Khả năng tích hợp với các hệ thống bảo mật khác, thiết bị mạng và phần mềm doanh nghiệp hiện có.
- Chi phí: So sánh về chi phí triển khai và duy trì giữa các giải pháp SIEM, từ đó xác định giải pháp phù hợp với ngân sách và quy mô của doanh nghiệp hoặc tổ chức.
- Hiệu quả: Đánh giá hiệu quả của mỗi hệ thống trong việc phát hiện các cuộc tấn công mạng, khả năng giảm thiểu cảnh báo giả (false positives) và mức độ dễ sử dụng đối với người quản trị.

b. Lựa chọn giải pháp SIEM:

Dựa trên phân tích các tiêu chí trên, đề tài sẽ đưa ra lựa chọn hệ thống SIEM phù hợp với môi trường triển khai thử nghiệm (môi trường học tập hoặc phòng thí nghiệm) và nghiên cứu chuyên sâu.

c. Phân tích và mô phỏng tấn công mạng:

- Triển khai các phương thức tấn công mạng phổ biến (DoS,brute-force,…) trong môi trường mô phỏng để thu thập log và phân tích dữ liệu.
- Sử dụng hệ thống SIEM để phát hiện các tấn công này, phân tích các log được tạo ra từ các cuộc tấn công và đánh giá khả năng của SIEM trong việc cảnh báo và phản hồi sự cố.
- Phân tích cách SIEM xử lý các loại tấn công khác nhau và đánh giá khả năng phân loại chính xác các sự kiện an ninh.

d. Thực hiện thử nghiệm triển khai:

Thử nghiệm triển khai giải pháp SIEM đã lựa chọn trong một môi trường mô phỏng hoặc phòng thí nghiệm học tập. Quá trình thử nghiệm bao gồm cấu hình hệ thống, tích hợp các nguồn log, và giám sát hoạt động hệ thống để đánh giá tính hiệu quả.

Thực hiện các cuộc tấn công mạng giả lập để kiểm tra tính năng của hệ thống SIEM trong phát hiện và ứng phó với sự cố.

5. Ý nghĩa của đề tài

a. Đối với doanh nghiệp và tổ chức

- Hệ thống SIEM đóng vai trò quan trọng trong việc nâng cao năng lực giám sát an ninh mạng, giúp phát hiện và phân tích các hoạt động bất thường và các cuộc tấn công mạng ngay từ giai đoạn đầu.
- SIEM cung cấp cho doanh nghiệp khả năng ứng phó nhanh chóng, giảm thiểu thiệt hại từ các cuộc tấn công mạng, bảo vệ dữ liệu và thông tin quan trọng khỏi bị xâm nhập hoặc mất mát.

Đề tài không chỉ góp phần nâng cao nhận thức về tầm quan trọng của SIEM mà còn cung cấp cái nhìn toàn diện về cách thức triển khai hệ thống SIEM một cách hiệu quả, từ việc lựa chọn giải pháp đến cấu hình và giám sát hệ thống.

b. Đối với nghiên cứu học thuật

Đề tài sẽ là tài liệu tham khảo hữu ích cho các sinh viên, nghiên cứu sinh và giảng viên trong lĩnh vực an ninh mạng, đặc biệt trong bối cảnh học tập và phòng thí nghiệm.

Nó cung cấp cơ hội cho sinh viên thực hành và hiểu rõ hơn về quy trình phát hiện tấn công mạng thông qua SIEM, từ việc mô phỏng tấn công đến việc sử dụng các công cụ bảo mật tiên tiến trong việc bảo vệ hệ thống mạng.

CHƯƠNG 1 : TỔNG QUAN VỀ TẤN CÔNG WEB

1.1 Giới thiệu về OWASP

1.1.1 Định nghĩa và mục tiêu

OWASP, viết tắt của *Open Web Application Security Project*, là một tổ chức phi lợi nhuận quốc tế chuyên về bảo mật ứng dụng web. OWASP hoạt động như một cộng đồng trực tuyến toàn cầu, cung cấp các bài viết, phương pháp, tài liệu, công cụ và công nghệ nhằm nâng cao nhận thức và cải thiện bảo mật cho ứng dụng web.

Một trong những nguyên tắc cốt lõi của OWASP là tất cả các tài liệu và công cụ mà họ phát triển đều được cung cấp miễn phí và dễ dàng truy cập trên trang web của họ. Điều này giúp mọi cá nhân và tổ chức có thể tiếp cận kiến thức và công nghệ để nâng cao bảo mật cho hệ thống của mình.

Tài liệu của OWASP bao gồm hướng dẫn, công cụ kiểm tra bảo mật, video hướng dẫn và các diễn đàn thảo luận. Trong đó, dự án nổi bật nhất và được biết đến rộng rãi là ***OWASP Top 10*** – một bảng xếp hạng 10 lỗ hổng bảo mật phổ biến nhất trong các ứng dụng web, được cập nhật định kỳ nhằm phản ánh các mối đe dọa bảo mật mới nhất.

1.1.2 Lịch sử phát triển

OWASP được thành lập vào năm 2001 bởi Mark Curphey. Qua thời gian, OWASP đã phát triển thành một tổ chức toàn cầu với cộng đồng đông đảo, bao gồm các nhà phát triển, chuyên gia bảo mật và doanh nghiệp. Một trong những tài nguyên nổi bật nhất của OWASP là dự án OWASP Top 10, ra mắt lần đầu vào năm 2003. Đây là bảng xếp hạng các lỗ hổng bảo mật phổ biến nhất trong các ứng dụng web, được cập nhật định kỳ.

Các điểm phát triển quan trọng trong lịch sử OWASP:

Năm 2001: OWASP được thành lập và nhanh chóng thu hút sự quan tâm của các chuyên gia bảo mật.

Năm 2003: Dự án OWASP Top 10 ra mắt, trở thành tài nguyên nổi tiếng nhất của OWASP, cung cấp bảng xếp hạng các lỗ hổng bảo mật phổ biến trong ứng dụng web, như Injection, XSS, và CSRF. OWASP Top 10 được cập nhật thường xuyên để phản ánh các nguy cơ mới nhất.

Năm 2005: OWASP phát hành OWASP Testing Guide, một tài liệu quan trọng hướng dẫn các nhà phát triển và kiểm thử viên phát hiện và khắc phục lỗ hổng trong các ứng dụng.

Năm 2007: Ra mắt dự án OWASP ZAP (Zed Attack Proxy), một công cụ miễn phí và mã nguồn mở cho việc kiểm tra bảo mật ứng dụng web.

Năm 2008: Dự án OWASP SAMM (Software Assurance Maturity Model) ra đời, cung cấp một mô hình để đánh giá và cải thiện quy trình phát triển phần mềm an toàn trong các tổ chức.

Năm 2010s: OWASP mở rộng mạnh mẽ với hàng nghìn thành viên và chương trình khu vực trên khắp thế giới, tổ chức các hội nghị như OWASP AppSec và cung cấp nhiều công cụ, dự án mã nguồn mở như Dependency-Check, OWASP ASVS (Application Security Verification Standard).

Hiện tại: OWASP tiếp tục đóng vai trò quan trọng trong việc định hình tiêu chuẩn bảo mật phần mềm, với hơn 200 dự án bảo mật mã nguồn mở, tài liệu giáo dục, và cộng đồng toàn cầu với các buổi hội thảo, sự kiện và cộng tác đa ngành.

1.2 Khái niệm Top 10 OWASP

1.2.1. Định nghĩa và vai trò

1.2.1.1 Định nghĩa

OWASP Top 10 không chỉ là một danh sách đơn thuần về các lỗ hổng bảo mật, mà còn là một hướng dẫn tiêu chuẩn nhằm giúp các tổ chức:

- Xác định và vá các lỗ hổng bảo mật trong các ứng dụng web.
- Nâng cao nhận thức về bảo mật trong các quy trình phát triển phần mềm.
- Cải thiện chất lượng bảo mật của ứng dụng ngay từ khâu thiết kế và phát triển.
- Xây dựng quy trình kiểm thử bảo mật toàn diện, giúp phát hiện các lỗ hổng trước khi chúng bị khai thác.

Tài liệu này được cập nhật định kỳ (*thường là mỗi 2-3 năm*) để phản ánh những thay đổi và sự phát triển trong kỹ thuật tấn công và bảo mật, qua đó cung cấp cho các tổ chức thông tin mới nhất về các lỗ hổng có nguy cơ cao nhất.

1.2.1.2 Vai trò

OWASP Top 10 đóng vai trò quan trọng trong việc bảo mật ứng dụng web. Dưới đây là các vai trò chính của nó:

- Công cụ giáo dục và nâng cao nhận thức: OWASP Top 10 là tài liệu quan trọng giúp cả người mới và chuyên gia nắm bắt các mối đe dọa bảo mật phổ biến. Nó là nền tảng để đào tạo các nhà phát triển và đội ngũ kỹ thuật về các nguy cơ bảo mật và cách phòng ngừa chúng.
- Hướng dẫn chiến lược bảo mật: OWASP Top 10 cung cấp tiêu chuẩn tối thiểu về bảo mật, giúp các tổ chức thiết lập chiến lược bảo mật ứng dụng web. Nó giúp xác định các khu vực cần chú trọng và phân bổ nguồn lực hợp lý để giảm thiểu rủi ro.
- Tiêu chuẩn hóa quy trình phát triển phần mềm an toàn: OWASP Top 10 giúp các tổ chức thiết kế và xây dựng quy trình phát triển phần mềm (SDLC) có bảo mật tích hợp ngay từ đầu, giảm thiểu chi phí và rủi ro khi phải sửa chữa các lỗ hổng sau này.
- Hỗ trợ tuân thủ các quy định bảo mật: OWASP Top 10 cung cấp hướng dẫn cụ thể giúp tổ chức tuân thủ các tiêu chuẩn bảo mật như GDPR và PCI DSS, đặc biệt trong việc bảo vệ thông tin cá nhân và dữ liệu nhạy cảm.
- Tăng cường bảo mật trong quy trình kiểm thử ứng dụng: OWASP Top 10 là cơ sở để xây dựng các bài kiểm thử bảo mật, giúp phát hiện và vá các lỗ hổng trước khi ứng dụng được triển khai, đảm bảo an toàn trước các cuộc tấn công tiềm ẩn.

1.2.2 Quy trình cập nhật

Quy trình cập nhật OWASP Top 10 là một quy trình chặt chẽ và toàn diện, được thực hiện với sự tham gia của nhiều nguồn dữ liệu từ khắp nơi trên thế giới. Nhờ vào việc liên tục cập nhật và phản ánh những thay đổi trong bối cảnh an ninh mạng, OWASP Top 10 luôn giữ được vai trò là tài liệu tham khảo quan trọng cho các tổ chức và cá nhân trong việc bảo mật ứng dụng web.Dưới đây là các bước :

- Thu thập dữ liệu: Dữ liệu được thu thập từ nhiều nguồn như các công ty bảo mật, nhà cung cấp phần mềm, chuyên gia bảo mật và các tổ chức như NIST,

CERT. Dữ liệu bao gồm các sự cố bảo mật, phương thức tấn công và tác động của chúng.

- Phân tích và phân loại: Dữ liệu được phân tích dựa trên tính phổ biến, mức độ nghiêm trọng và độ phức tạp trong việc khai thác. Từ đó, các lỗ hổng được phân loại để dễ dàng đánh giá.
- Đánh giá mức độ nghiêm trọng: Các lỗ hổng được đánh giá dựa trên tác động tiềm tàng, tần suất xuất hiện và khả năng phát hiện. Những lỗ hổng nghiêm trọng và phổ biến nhất sẽ được đưa vào danh sách.
- Cập nhật danh sách và hướng dẫn phòng chống: Danh sách OWASP Top 10 mới được cập nhật kèm theo hướng dẫn về nguyên nhân, tác động, và cách phòng chống cụ thể cho từng lỗ hổng.
- Phản hồi từ cộng đồng: OWASP nhận phản hồi và đóng góp từ cộng đồng bảo mật toàn cầu để đảm bảo danh sách luôn phản ánh chính xác mối đe dọa hiện tại.

1.3 Sơ lược các lỗ hổng trong bảng TOP10 OWASP những năm gần đây

1.3.1 Tổng hợp các lỗ hổng OWASP Top 10 năm 2017

Phiên bản OWASP Top 10 năm 2017 liệt kê các lỗ hổng phổ biến nhất tại thời điểm đó, bao gồm các lỗ hổng truyền thống và một số mới nổi, như sau:

Bảng 1: TOP10 OWASP năm 2017

STT	Tên lỗ hổng	Mô tả	Tác động
1	Injection	Đây là lỗ hổng cho phép kẻ tấn công chèn các mã độc vào các truy vấn như SQL, NoSQL, OS, và LDAP thông qua đầu vào của người dùng không được kiểm soát.	Kẻ tấn công có thể thay đổi câu truy vấn hoặc lệnh hệ thống, dẫn đến rò rỉ thông tin, mất dữ liệu, hoặc kiểm soát hệ thống từ xa.
2	Broken Authentication	Các cơ chế xác thực không an toàn hoặc dễ bị	Kẻ tấn công có thể giả mạo danh tính người dùng,

	(Xác thực không an toàn)	khai thác, chẳng hạn như mật khẩu yếu hoặc cơ chế phiên không an toàn.	chiếm quyền kiểm soát tài khoản hoặc hệ thống.
3	Sensitive Data Exposure (Rò rỉ dữ liệu nhạy cảm)	Dữ liệu nhạy cảm không được bảo vệ đúng cách (mã hóa yếu hoặc không đủ kiểm soát truy cập).	Kẻ tấn công có thể truy cập dữ liệu nhạy cảm như thông tin thẻ tín dụng hoặc thông tin nhận dạng cá nhân.
4	XML External Entities (XXE)	Lỗ hổng do cấu hình không an toàn trong phân tích XML, cho phép kẻ tấn công truy cập dữ liệu nhạy cảm hoặc thực hiện các cuộc tấn công như DoS hoặc SSRF.	Kẻ tấn công có thể thực hiện đọc file trên máy chủ hoặc tương tác với các hệ thống nội bộ không được phép.
5	Broken Access Control (Kiểm soát truy cập không an toàn)	Các quy tắc kiểm soát truy cập không được thực hiện chính xác, cho phép người dùng truy cập vào các tài nguyên mà họ không được phép.	Kẻ tấn công có thể truy cập và chỉnh sửa dữ liệu hoặc các chức năng của hệ thống mà không được cấp quyền.
6	Security Misconfiguration (Cấu hình bảo mật không an toàn)	Cấu hình bảo mật không chính xác hoặc thiếu cẩn trọng, ví dụ như để lộ thông tin hệ thống, bật các tính năng không cần thiết.	Tạo điều kiện cho kẻ tấn công khai thác các lỗ hổng tiềm tàng trong hệ thống.
7	Cross-Site Scripting (XSS)	Kẻ tấn công có thể chèn mã độc vào các trang web, lừa người dùng thực thi các đoạn mã nguy	Kẻ tấn công có thể đánh cắp phiên làm việc của người dùng, chiếm quyền điều khiển hoặc thực hiện

		hiểm (JavaScript, HTML).	các hoạt động nguy hại khác.
8	Insecure Deserialization (Deserialization không an toàn)	Lỗ hổng xảy ra khi dữ liệu không an toàn được chuyển đổi thành đối tượng, cho phép kẻ tấn công thực thi mã hoặc kiểm soát hệ thống.	Kẻ tấn công có thể lợi dụng để kiểm soát ứng dụng hoặc thực thi mã độc.
9	Using Components with Known Vulnerabilities (Sử dụng thành phần có lỗ hổng bảo mật)	Sử dụng các thư viện, framework hoặc thành phần đã có lỗ hổng bảo mật được công bố.	Kẻ tấn công có thể lợi dụng các lỗ hổng này để tấn công hệ thống.
10	Insufficient Logging & Monitoring (Thiếu giám sát và ghi nhật ký)	Hệ thống thiếu giám sát và ghi nhật ký đầy đủ, gây khó khăn trong việc phát hiện các cuộc tấn công và ứng phó kịp thời.	Kẻ tấn công có thể xâm nhập và duy trì sự hiện diện lâu dài mà không bị phát hiện.

1.3.2 Tổng hợp các lỗ hổng TOP10 OWASP năm 2021

Phiên bản OWASP Top 10 năm 2021 cập nhật danh sách để phản ánh những xu hướng tấn công mới, với một số thay đổi quan trọng:

Bảng 2: TOP10 OWASP 2021

STT	Tên lỗ hổng	Mô tả	Tác động

1	**Broken Access Control**	Lỗ hổng trong kiểm soát truy cập đã tăng lên vị trí đầu tiên do ngày càng phổ biến.	Kẻ tấn công có thể thực hiện hành động trái phép, chẳng hạn như sửa hoặc xóa dữ liệu, trên các tài nguyên mà họ không được phép.
2	**Cryptographic Failures (Thất bại trong mã hóa)**	Đổi tên từ "Sensitive Data Exposure", tập trung vào các vấn đề bảo mật liên quan đến mã hóa và quản lý khóa.	Dữ liệu nhạy cảm có thể bị lộ nếu cơ chế mã hóa không đủ an toàn.
3	**Injection**	Dù giảm bậc nhưng vẫn là một mối đe dọa nghiêm trọng, bao gồm các hình thức chèn SQL, NoSQL, và OS.	Kẻ tấn công có thể thao túng cơ sở dữ liệu hoặc thực hiện lệnh hệ thống trái phép.
4	**Insecure Design**	**Lỗ hổng mới** này nhấn mạnh vào việc thiết kế ứng dụng không an toàn từ ban đầu.	Các hệ thống có thiết kế không bảo mật dễ bị khai thác ngay cả khi các biện pháp bảo vệ khác được thực thi.
5	**Security Misconfiguration**	Cấu hình không chính xác tiếp tục là một lỗ hổng phổ biến, đặc biệt trong các hệ thống phức tạp với nhiều thành phần.	Kẻ tấn công có thể khai thác lỗ hổng bảo mật không được bảo vệ đúng cách do cấu hình kém.

6	**Vulnerable and Outdated Components**	"**Using Components with Known Vulnerabilities"** đã cải tiến**,** lỗ hổng này tập trung vào các thành phần không cập nhật.	Dùng các thành phần đã lỗi thời có thể tạo điều kiện cho các cuộc tấn công.
7	**Identification and Authentication Failures**	"Broken Authentication", đã bị thay thế ,mở rộng hơn về các lỗi liên quan đến nhận dạng và xác thực.	Kẻ tấn công có thể chiếm quyền tài khoản, giả mạo danh tính.
8	**Software and Data Integrity Failures**	**Lỗ hổng mới** nhắm vào các tấn công vào tính toàn vẹn của phần mềm, chẳng hạn như việc sử dụng các cập nhật không an toàn hoặc các nguồn không đáng tin cậy.	Kẻ tấn công có thể thay đổi hoặc kiểm soát các thành phần phần mềm.
9	**Security Logging and Monitoring Failures**	Tập trung hơn vào khả năng phát hiện và phản ứng với các cuộc tấn công.	Việc thiếu giám sát làm cho hệ thống dễ bị tấn công mà không bị phát hiện.
10	**Server-Side Request Forgery (SSRF)**	**Lỗ hổng mới,** xuất hiện khi máy chủ xử lý các yêu cầu không tin cậy từ phía máy khách, dẫn đến truy cập trái phép vào các hệ thống nội bộ.	Kẻ tấn công có thể khai thác hệ thống nội bộ hoặc thậm chí các dịch vụ đám mây thông qua yêu cầu phía máy chủ.

1.4 Tầm quan trọng của OWASP Top 10 trong phát triển phần mềm

a. Hỗ trợ nhận thức về bảo mật

OWASP Top 10 cung cấp một danh sách các lỗ hổng bảo mật phổ biến nhất, giúp các nhà phát triển nhận diện các mối đe dọa tiềm ẩn cho ứng dụng của họ và nhấn mạnh tầm quan trọng của việc tích hợp bảo mật từ giai đoạn đầu trong quy trình phát triển.

b. Điều chỉnh quy trình phát triển phần mềm

Danh sách này giúp các nhà phát triển điều chỉnh quy trình phát triển để giảm thiểu rủi ro bảo mật thông qua:

- Thiết kế an toàn: Áp dụng các nguyên tắc thiết kế bảo mật.
- Kiểm tra bảo mật: Tích hợp các bước kiểm tra vào từng giai đoạn phát triển, sử dụng công cụ tự động để phát hiện lỗ hổng.
- Đào tạo: Nâng cao kỹ năng và kiến thức bảo mật thông qua các khóa học và hội thảo.

c. Khung tham chiếu cho kiểm tra bảo mật

OWASP Top 10 cung cấp khung tham chiếu cho việc kiểm tra bảo mật trong các giai đoạn phát triển phần mềm:

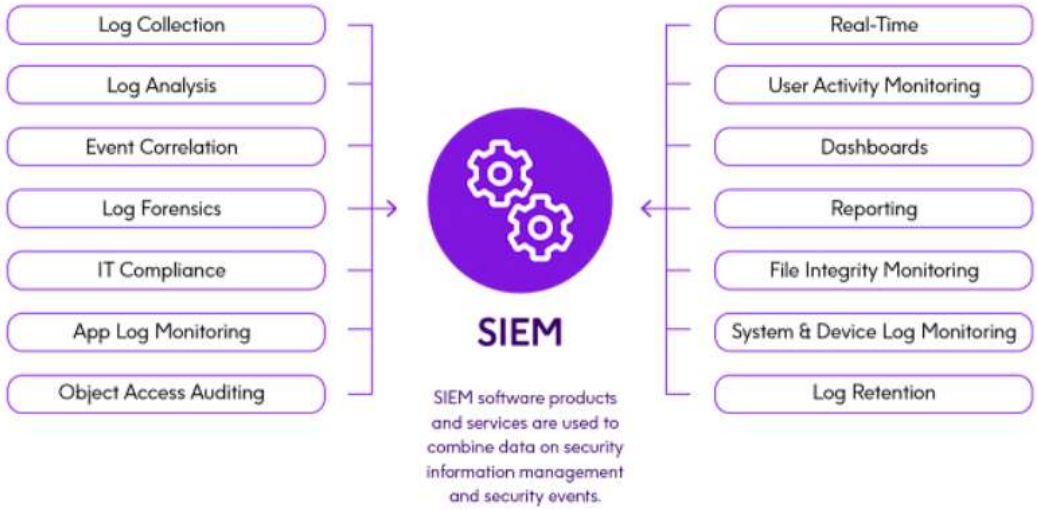
- Thiết kế: Tích hợp yêu cầu bảo mật vào thiết kế hệ thống.
- Viết mã: Sử dụng phương pháp lập trình an toàn.
- Kiểm thử: Thực hiện kiểm tra bảo mật trước khi triển khai.
- Triển khai: Đảm bảo cấu hình và biện pháp bảo mật được thực hiện đúng cách.

CHƯƠNG 2. TỔNG QUAN VỀ HỆ THỐNG GIÁM SÁT SIEM

2.1 Khái niệm

Hệ thống SIEM là viết tắt của cụm từ tiếng Anh Security Infomation and Event Management. Đây là một lĩnh vực trong bảo mật máy tính, nơi các sản phẩm và dịch vụ phần mềm kết hợp quản lý thông tin bảo mật (SIM) và quản lý sự kiện bảo mật (SEM). Các giải pháp SIEM có khả năng phát hiện các mối đe dọa và lỗ hổng bảo mật tiềm ẩn bằng cách phân tích dữ liệu nhật ký và sự kiện trong thời gian thực từ nhiều nguồn khác nhau bao gồm cả mạng, bảo mật, máy chủ, cơ sở dữ liệu và ứng dụng.

Họ cung cấp các cảnh báo và thông báo cho các nhóm bảo mật để điều tra và phản hồi thêm. Các công cụ SIEM đã phát triển đáng kể trong những năm qua và hiện được coi là một công cụ thiết yếu để phát hiện, điều tra và ứng phó với các mối đe dọa an ninh mạng nâng cao.



Hình 2.1: Các thành phần chính của SIEM

Ban đầu, các nền tảng SIEM được phát triển dưới dạng công cụ quản lý log. Chúng kết hợp hai chức năng chính là Security Information Management (SIM) và Security Event Management (SEM), cho phép giám sát và phân tích theo thời gian thực các sự kiện liên quan đến bảo mật.

Ngoài ra, các nền tảng này cũng hỗ trợ việc theo dõi và ghi log dữ liệu bảo mật để đáp ứng yêu cầu compliance hoặc kiểm toán. Gartner đã đặt ra thuật ngữ SIEM vào năm 2005 khi kết hợp hai công nghệ SIM và SEM.

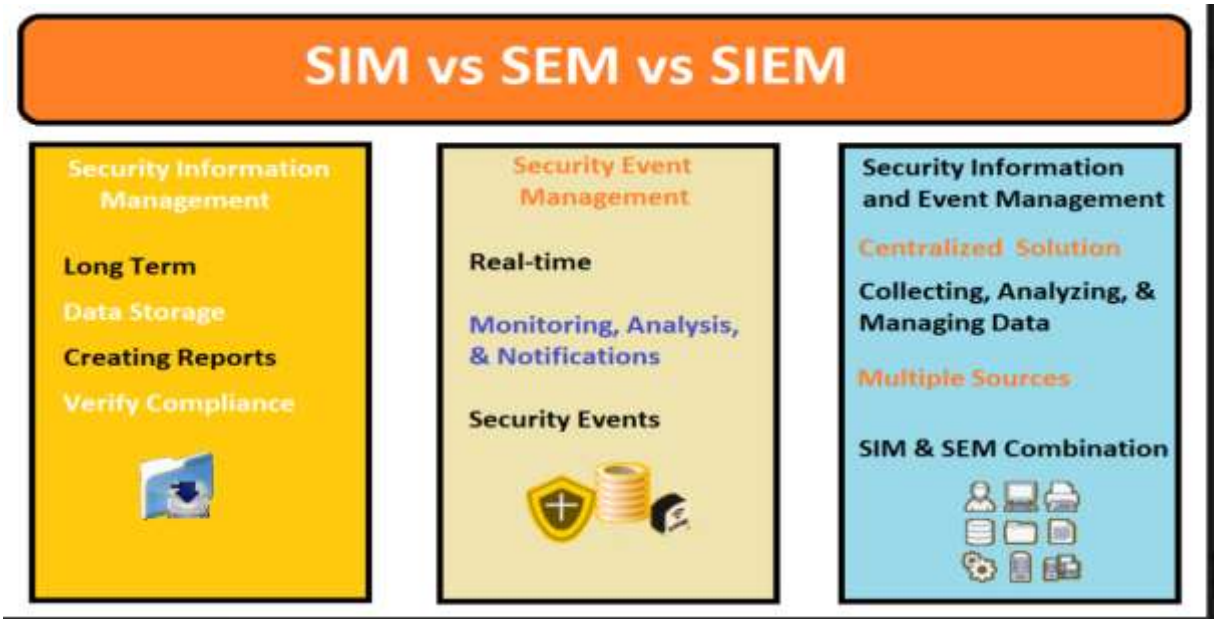
Qua nhiều năm phát triển, phần mềm SIEM đã tích hợp thêm các công nghệ phân tích bảo mật tiên tiến như User and Entity Behavior Analytics (UEBA), AI, và machine learning để nhận diện các hành vi bất thường và các chỉ số của mối đe dọa tinh vi. Ngày nay, SIEM đã trở thành một thành phần quan trọng trong các Security Operation Center (SOC) hiện đại, chủ yếu phục vụ cho việc giám sát bảo mật và quản lý tuân thủ.

SIEM hoạt động như một kho lưu trữ trung tâm do nhật ký hệ thống tạo ra, thông qua các quy tắc logic mà bạn thiết lập, để chọn ra các sự kiện quan tâm cụ thể. Từ SIEM, khi có thể bị tấn công mạng ta có thể xem được nhiều loại thông tin và các sự kiện từ nhiều thiết bi.

Tóm lại, SIEM cung cấp cho các tổ chức khả năng quan sát hoạt động trong mạng của họ để họ có thể ứng phó nhanh chóng với các cuộc tấn công qua mạng tiềm ẩn và đáp ứng các yêu cầu tuân thủ.

2.2 Cách thức hoạt động của hệ thống giám sát

Giải pháp SIEM thu thập dữ liệu nhật ký và sự kiện từ các thành phần khác nhau của mạng lưới doanh nghiệp. Sau khi chuẩn hóa dữ liệu, nó sử dụng thông tin về mối đe dọa, các quy tắc tích hợp sẵn, và các chức năng phân tích nâng cao để phát hiện các sự cố bảo mật theo thời gian thực. Nói cách khác, SIEM cung cấp một cái nhìn toàn diện về bảo mật thông tin của tổ chức. Tùy thuộc vào kiến trúc của nó, SIEM sắp xếp các cảnh báo vào các danh mục khác nhau như phần mềm độc hại, đăng nhập thất bại, đăng nhập thành công, các hoạt động có thể gây hại khác, v.v.



Hình 2.2: SIM, SEM và SIEM

SIEM kết hợp hai công nghệ: ***Security Information Management (SIM)*** và ***Security Event Managementt (SEM).*** Trong các giải pháp SIEM hiện đại, rất khó để tách biệt hai thành phần này. SIM chủ yếu lo việc thu thập dữ liệu từ các nguồn nhật ký và tạo ra các báo cáo mong muốn. Trong khi đó, SEM thực hiện việc giám sát theo thời gian thực các hệ thống của doanh nghiệp để phát hiện mối đe dọa và liên kết các sự kiện.

Khi giải pháp SIEM xác định được một mối đe dọa tiềm ẩn, nó tạo ra các cảnh báo để thông báo cho đội ngũ bảo mật. Dựa trên các quy tắc đã định trước, mức độ ưu tiên của cảnh báo có thể là thấp, trung bình hoặc cao. Ví dụ, nếu tài khoản người dùng X tạo ra mười lần đăng nhập trong năm phút, điều này có thể được coi là hoạt động đáng ngờ. Tuy nhiên, rất có thể người dùng X đã quên mật khẩu và không thể đăng nhập. Giả sử cùng một tài khoản người dùng trải qua 200 lần đăng nhập trong cùng khoảng thời gian đó, giải pháp SIEM sẽ gán hoạt động này là một sự cố nghiêm trọng vì có thể đó là một cuộc tấn công bẻ khóa mật khẩu.

2.2.1. Thành phần của SIEM

Log Collection (Thu thập nhật ký)

Đây là bước đầu tiên trong quy trình SIEM, nơi dữ liệu được thu thập từ nhiều nguồn khác nhau trong hệ thống.Trong một hệ thống SIEM (Security Information and Event Management), dữ liệu được thu thập từ nhiều nguồn khác nhau nhằm đảm bảo tính toàn diện.

a. Nguồn dữ liệu

- ❖ Thiết bị mạng
  - Router và Switch: Nhật ký từ các router và switch cung cấp thông tin về lưu lượng mạng, cấu hình, và các thay đổi liên quan đến tuyến mạng.
  - Tường lửa: Nhật ký truy cập mạng, các chính sách chặn/lọc, và các sự kiện liên quan đến tấn công mạng (như DoS, port scanning).
  - IDS/IPS: Cung cấp dữ liệu về các dấu hiệu xâm nhập, cảnh báo về các mẫu tấn công hoặc các vi phạm chính sách bảo mật.
- ❖ Máy chủ
  - Máy chủ ứng dụng: Theo dõi các yêu cầu HTTP, HTTPS, và giao dịch giữa các thành phần ứng dụng.

- Máy chủ cơ sở dữ liệu: Nhật ký truy vấn, thay đổi dữ liệu, và truy cập người dùng.
- Máy chủ web: Ghi nhận lưu lượng truy cập, lỗi, hoặc các hành vi không hợp lệ, như tấn công SQL injection hay cross-site scripting (XSS).

❖ Ứng dụng

- Ứng dụng doanh nghiệp: Bao gồm các hệ thống ERP, CRM, hoặc các ứng dụng tùy chỉnh. Cung cấp dữ liệu về hoạt động giao dịch, lỗi logic, hoặc vi phạm chính sách sử dụng.
- Hệ thống quản lý nội dung (CMS): Theo dõi đăng nhập, thay đổi dữ liệu, và các cuộc tấn công qua plugin, giao diện API.
- Phần mềm tùy chỉnh: Dữ liệu được thiết kế riêng cho tổ chức, thường bao gồm nhật ký giao tiếp API, lỗi, hoặc hoạt động của người dùng.

❖ Thiết bị đầu cuối

- Máy tính cá nhân: Thông tin về quyền truy cập, hành vi người dùng, và các phần mềm được sử dụng.
- Điện thoại thông minh: Dữ liệu liên quan đến vị trí, ứng dụng đã cài đặt, và hoạt động mạng di động.
- Thiết bị IoT: Nhật ký trạng thái thiết bị, hoạt động mạng, và các lỗ hổng có thể khai thác.

b. Phương pháp thu thập dữ liệu

Để đảm bảo việc thu thập thông tin hiệu quả và đồng nhất từ các nguồn, SIEM sử dụng các kỹ thuật và giao thức thu thập khác nhau:

❖ Giao thức Syslog

- Gửi nhật ký từ các thiết bị mạng và máy chủ đến hệ thống SIEM.
- Dễ dàng triển khai với hầu hết các thiết bị hỗ trợ chuẩn RFC 5424.
- Hữu ích trong việc thu thập nhật ký thời gian thực từ tường lửa, router, và switch.

❖ SNMP (Simple Network Management Protocol)

- Thu thập thông tin trạng thái, cấu hình, và cảnh báo từ thiết bị mạng.
- Hỗ trợ việc giám sát hiệu năng và phát hiện các bất thường như băng thông tăng đột biến.

❖ API

- Ứng dụng đám mây: Các dịch vụ như AWS, Azure, hoặc Google Cloud thường cung cấp API để truy xuất dữ liệu nhật ký và sự kiện bảo mật.
- Ứng dụng doanh nghiệp: API giúp truy cập dữ liệu từ các hệ thống nội bộ một cách linh hoạt.

❖ Agent (Tác nhân)

- Phần mềm nhỏ gọn được cài đặt trên thiết bị đầu cuối, máy chủ, hoặc ứng dụng.
- Thu thập dữ liệu cục bộ và gửi trực tiếp đến SIEM theo thời gian thực.
- Ưu điểm: Có thể thu thập thông tin chi tiết như quá trình chạy, thay đổi tệp, và các hoạt động truy cập.

c. Lợi ích của việc thu thập dữ liệu hiệu quả

❖ Kho dữ liệu phong phú

- Tích hợp đa nguồn cho phép phân tích tổng quan toàn bộ hệ thống, hỗ trợ xây dựng một cơ sở dữ liệu trung tâm.
- Tạo điều kiện cho việc áp dụng các công nghệ học máy nhằm phát hiện các mẫu bất thường hoặc các dấu hiệu tấn công tiềm ẩn.

❖ Phát hiện sự cố và hành vi bất thường sớm

- Việc thu thập dữ liệu thời gian thực từ tất cả các nguồn giúp phát hiện và phản ứng nhanh với các sự cố.
- Hỗ trợ giám sát các chỉ số chính (KPI), phát hiện các thay đổi cấu hình trái phép hoặc hành vi xâm nhập.

❖ Tăng cường khả năng giám sát và phân tích

- SIEM có thể sử dụng dữ liệu để tái hiện lại dòng sự kiện trong một cuộc điều tra bảo mật.

- Hỗ trợ phát hiện các cuộc tấn công phức tạp như APT (Advanced Persistent Threat).

❖ Tuân thủ và kiểm toán

- Dữ liệu thu thập được cung cấp bằng chứng cho việc tuân thủ các tiêu chuẩn như GDPR, ISO 27001, hoặc PCI DSS.
- Dễ dàng thực hiện kiểm tra bảo mật định kỳ và cung cấp thông tin chi tiết về hoạt động trong hệ thống.

Log Analysis (Phân tích nhật ký)

Phân tích nhật ký là bước tiếp theo sau khi dữ liệu nhật ký được thu thập. Đây là giai đoạn trung tâm của hệ thống SIEM, nơi các bản ghi được kiểm tra và phân tích để phát hiện các hành vi bất thường hoặc dấu hiệu của các cuộc tấn công. Phân tích nhật ký không chỉ tập trung vào việc tìm kiếm các sự kiện cụ thể mà còn khám phá các mẫu và xu hướng có thể cho thấy nguy cơ tiềm ẩn.

a. Công cụ và phương pháp phân tích nhật ký

❖ Phân tích theo quy tắc (Rule-Based Analysis): Sử dụng các quy tắc đã được thiết lập từ trước để xác định các hành vi đáng ngờ dựa trên các điều kiện cụ thể. Đây là cách tiếp cận dựa trên logic, thường được triển khai dễ dàng và hiệu quả cho các vấn đề đã biết.

Ví dụ:Nhiều lần đăng nhập thất bại liên tiếp từ cùng một địa chỉ IP trong thời gian ngắn có thể là dấu hiệu của tấn công Brute Force.Đột nhiên có nhiều truy cập từ một tài khoản vào các tệp tin nhạy cảm.

❖ Ưu điểm:

- Đơn giản, dễ triển khai.
- Hiệu quả với các sự kiện đã biết.

❖ Hạn chế:

- Phụ thuộc vào việc xây dựng và bảo trì quy tắc.
- Khó phát hiện các mối đe dọa mới hoặc phức tạp.

b. Phân tích dựa trên học máy (Machine Learning Analysis)

Sử dụng các thuật toán học máy để tự động phân tích dữ liệu nhật ký, phát hiện các hành vi bất thường và xác định các mối đe dọa chưa từng biết trước đó.

❖ Ứng dụng:

- Phân cụm (Clustering): Sử dụng thuật toán như K-means hoặc Isolation Forest để xác định các cụm hành vi bình thường và bất thường.
- Phân loại (Classification): Sử dụng các thuật toán như Random Forest hoặc Support Vector Machine (SVM) để phân loại sự kiện là bình thường hoặc nguy cơ.

❖ Ưu điểm:

- Hiệu quả trong việc phát hiện các tấn công chưa được biết đến.
- Tự động cải thiện qua thời gian nếu được huấn luyện liên tục.

❖ Hạn chế:

- Yêu cầu nhiều dữ liệu chất lượng cao.
- Cần tài nguyên tính toán lớn.

c. Phân tích hành vi người dùng (UBA - User Behavior Analytics)

Theo dõi và phân tích hành vi của người dùng theo thời gian để phát hiện các thay đổi bất thường, từ đó phát hiện các mối đe dọa tiềm ẩn.

❖ Ứng dụng:

- Phát hiện người dùng bị xâm nhập khi tài khoản của họ thực hiện các hành vi bất thường (như đăng nhập vào lúc nửa đêm hoặc truy cập vào hệ thống không liên quan).
- Xác định các tài khoản có nguy cơ cao dựa trên mẫu hoạt động.

❖ Ưu điểm:

- Hiệu quả trong việc phát hiện các mối đe dọa nội bộ (insider threats).
- Tập trung vào hành vi thay vì chỉ dựa vào quy tắc cố định.

❖ Hạn chế: Khó thiết lập ngưỡng hành vi bình thường cho từng người dùng.

d. Lợi ích của phân tích nhật ký

❖ Phát hiện nhanh chóng các mối đe dọa tiềm tàng

- Phân tích nhật ký giúp tự động hóa việc giám sát, phát hiện các bất thường trong thời gian thực, giảm thiểu thời gian phát hiện và phản ứng.
- Các tấn công như Ransomware, Phishing, hoặc APT (Advanced Persistent Threat) có thể được phát hiện sớm.

❖ Cung cấp thông tin chi tiết về cách thức tấn công

- Phân tích nhật ký giúp hiểu rõ cách thức tấn công diễn ra bằng cách tái tạo lại dòng sự kiện, cung cấp dữ liệu quan trọng cho đội ngũ bảo mật.
- Hỗ trợ điều tra pháp lý khi cần thiết.

❖ Cải thiện khả năng phòng thủ của hệ thống

- Nhờ phân tích sâu, các lỗ hổng trong hệ thống có thể được xác định, từ đó nâng cấp các biện pháp bảo vệ.

❖ Hỗ trợ tuân thủ quy định: Giúp tổ chức đáp ứng các tiêu chuẩn bảo mật như GDPR, PCI DSS, hoặc ISO 27001 thông qua việc lưu trữ và phân tích nhật ký có hệ thống.

e. Thách thức trong phân tích nhật ký

❖ Khối lượng dữ liệu lớn: Hệ thống phải xử lý hàng triệu bản ghi mỗi ngày, dẫn đến khó khăn trong việc phân tích dữ liệu một cách nhanh chóng và chính xác.

❖ Đa dạng nguồn dữ liệu: Dữ liệu nhật ký đến từ nhiều nguồn với định dạng khác nhau, đòi hỏi phải chuẩn hóa trước khi phân tích.

❖ Độ phức tạp của các cuộc tấn công: Các tấn công hiện đại ngày càng tinh vi, khó phát hiện bằng các phương pháp truyền thống.

❖ Chi phí triển khai và duy trì: Yêu cầu tài nguyên lớn, cả về hạ tầng kỹ thuật lẫn đội ngũ nhân sự.

Event Correlation (Tương quan sự kiện)

Event Correlation (tương quan sự kiện) là một trong những chức năng quan trọng của hệ thống SIEM, giúp phân tích và kết hợp các sự kiện bảo mật từ nhiều nguồn khác nhau. Điều này giúp nhận diện các cuộc tấn công phức tạp, đặc biệt là những cuộc tấn công diễn ra qua nhiều giai đoạn hoặc có các yếu tố không rõ ràng.

a. Kỹ thuật tương quan sự kiện

- ❖ Tương quan theo thời gian (Temporal Correlation): Xác định các sự kiện xảy ra đồng thời hoặc trong một khoảng thời gian ngắn nhằm phát hiện ra các mối liên hệ.

Ví dụ: *Một tường lửa ghi nhận các kết nối bất thường vào lúc 2:00 sáng, và ngay sau đó một máy chủ bị ghi nhận tải xuống phần mềm độc hại.*

- ❖ Ứng dụng: Phát hiện các chuỗi tấn công hoặc các sự kiện có tính liên tục như brute-force kết hợp với exfiltration (rò rỉ dữ liệu).
- ❖ Tương quan theo điều kiện (Conditional Correlation): Sự kiện được phát hiện dựa trên các điều kiện cụ thể.

Ví dụ:

- *Một tài khoản người dùng bị khóa sau 5 lần đăng nhập thất bại liên tiếp từ các địa chỉ IP khác nhau.*
- *Một sự kiện truy cập xảy ra từ vị trí bất thường đối với tài khoản người dùng cụ thể.*

- ❖ Ứng dụng:Theo dõi các nỗ lực tấn công tài khoản hoặc các hành động vượt quá quyền hạn của người dùng.
- ❖ Tương quan thông minh (Intelligent Correlation): Áp dụng trí tuệ nhân tạo (AI) và học máy (Machine Learning) để phát hiện các mẫu hành vi bất thường, giúp tìm ra các mối đe dọa mới hoặc chưa biết.

Ví dụ: *Phát hiện các cuộc tấn công nâng cao (Advanced Persistent Threats - APT) diễn ra trong thời gian dài với các bước nhỏ.*

- ❖ Ứng dụng:Tự động phát hiện các cuộc tấn công chưa được ghi nhận trong cơ sở dữ liệu dấu hiệu (signature database).

b. Lợi ích của Event Correlation

- ❖ Phát hiện các cuộc tấn công phức tạp: Giúp phát hiện các mối đe dọa diễn ra theo nhiều giai đoạn, như tấn công phishing, privilege escalation, và sau đó là data exfiltration.

❖ Tăng cường khả năng phân tích và phản ứng: Kết hợp các thông tin từ nhiều nguồn để cung cấp cái nhìn tổng thể, từ đó đưa ra các hành động phản ứng kịp thời, chẳng hạn như chặn IP độc hại hoặc cô lập hệ thống bị tấn công.

Log Forensics (Pháp y nhật ký)

Log Forensics (pháp y nhật ký) là quy trình phân tích và khai thác dữ liệu từ các bản ghi nhật ký sau sự cố bảo mật. Mục tiêu chính là điều tra và xác định nguyên nhân, phạm vi, và phương thức của cuộc tấn công để hỗ trợ việc ngăn ngừa sự cố trong tương lai và cung cấp bằng chứng pháp lý nếu cần.

a. Quy trình điều tra pháp y nhật ký

❖ Phân tích chi tiết (Detailed Analysis)

Trích xuất và phân tích các bản ghi liên quan đến một sự kiện cụ thể để tìm kiếm thông tin quan trọng.

Ví dụ:Ghi nhận các địa chỉ IP, mã trạng thái HTTP, và thông tin đăng nhập không thành công trong một chuỗi tấn công brute-force.

❖ Ứng dụng: Xác định các giai đoạn của cuộc tấn công, từ việc dò quét lỗ hổng đến khai thác và thực thi mã độc.

❖ Truy vết nguồn gốc (Tracing the Source): Sử dụng dữ liệu nhật ký để xác định nguồn gốc của tấn công, như địa chỉ IP khởi tạo, thiết bị bị xâm nhập đầu tiên, hoặc tài khoản bị khai thác.

❖ Ứng dụng: Lần theo hành vi của hacker trong mạng nội bộ để phát hiện những máy bị lây nhiễm khác.

b. Khôi phục dữ liệu (Data Recovery)

Sử dụng thông tin từ nhật ký để xác định các dữ liệu bị mất hoặc bị thay đổi trong cuộc tấn công, đồng thời áp dụng các biện pháp để khôi phục chúng.

❖ Ứng dụng:

- Phục hồi dữ liệu cơ sở dữ liệu bị xóa.
- Xác định những tệp tin nào đã bị mã hóa trong cuộc tấn công ransomware.

c. Lợi ích của Log Forensics

- ❖ Cải thiện khả năng phòng ngừa sự cố: Thông qua việc phân tích chi tiết các sự cố trước đó, tổ chức có thể xây dựng các biện pháp phòng ngừa hiệu quả hơn để đối phó với các mối đe dọa tương tự trong tương lai.
- ❖ Hỗ trợ điều tra pháp lý: Nhật ký cung cấp các bằng chứng quan trọng cho quá trình điều tra pháp lý, đảm bảo rằng mọi hoạt động đáng ngờ hoặc vi phạm được ghi nhận đầy đủ và chính xác.
- ❖ Tăng cường khả năng ứng phó sự cố: Phân tích nhật ký nhanh chóng giúp xác định và ngăn chặn các cuộc tấn công đang diễn ra, giảm thiểu thiệt hại cho hệ thống.

IT Compliance (Tuân thủ IT)

Trong bối cảnh các quy định và tiêu chuẩn bảo mật ngày càng nghiêm ngặt, việc tuân thủ các quy định pháp lý và tiêu chuẩn bảo mật trở thành một yêu cầu thiết yếu cho tổ chức. Hệ thống SIEM hỗ trợ tự động hóa các quy trình liên quan đến tuân thủ, giúp tổ chức đáp ứng các yêu cầu từ các tiêu chuẩn quốc tế như GDPR (General Data Protection Regulation), ISO 27001, PCI DSS (Payment Card Industry Data Security Standard), và HIPAA (Health Insurance Portability and Accountability Act).

a. Công cụ hỗ trợ cho IT Compliance

- ❖ Báo cáo tuân thủ (Compliance Reporting): SIEM tạo ra các báo cáo tự động, định kỳ hoặc theo yêu cầu, phục vụ cho các cuộc kiểm tra tuân thủ. Các báo cáo này bao gồm chi tiết về nhật ký, sự kiện bảo mật, và bất kỳ vi phạm nào đối với chính sách nội bộ hoặc quy định bên ngoài.
- ❖ Ứng dụng:
  - Báo cáo sự kiện truy cập không hợp lệ vào dữ liệu nhạy cảm.
  - Theo dõi trạng thái tuân thủ của các hệ thống mạng và ứng dụng.
- ❖ Lợi ích:
  - Giảm thời gian chuẩn bị báo cáo thủ công.
  - Tăng độ chính xác của dữ liệu trình bày cho kiểm toán viên.

b. Lưu trữ nhật ký (Log Retention)

SIEM hỗ trợ lưu trữ dữ liệu nhật ký trong khoảng thời gian được yêu cầu bởi các quy định pháp lý hoặc chính sách của tổ chức. Điều này bao gồm việc lưu trữ lâu dài để phục vụ cho các mục đích điều tra hoặc kiểm toán trong tương lai.

❖ Ứng dụng:

- Lưu giữ nhật ký trong 7 năm để tuân thủ SOX (Sarbanes-Oxley Act).
- Lưu nhật ký truy cập bệnh nhân trong 5 năm theo tiêu chuẩn HIPAA.

❖ Lợi ích:

- Đảm bảo rằng dữ liệu được bảo vệ và truy xuất dễ dàng khi cần.
- Giảm thiểu rủi ro mất dữ liệu do không tuân thủ.

c. Lợi ích của IT Compliance

❖ Giảm thiểu rủi ro pháp lý và bảo vệ dữ liệu người dùng: Bằng cách tự động hóa và duy trì quy trình tuân thủ, tổ chức có thể tránh được các hình phạt pháp lý nghiêm trọng hoặc rủi ro kiện tụng liên quan đến bảo mật dữ liệu.

❖ Tăng cường niềm tin từ khách hàng và các bên liên quan: Tuân thủ các tiêu chuẩn bảo mật giúp nâng cao uy tín và củng cố niềm tin từ khách hàng, đối tác, và nhà đầu tư.

App Log Monitoring (Giám sát nhật ký ứng dụng)

Ứng dụng doanh nghiệp thường là mục tiêu tấn công chính của hacker, bởi chúng chứa dữ liệu nhạy cảm và là cửa ngõ chính để truy cập vào hệ thống nội bộ. SIEM cung cấp khả năng giám sát và phân tích các nhật ký liên quan đến ứng dụng nhằm phát hiện sớm các lỗ hổng bảo mật hoặc dấu hiệu tấn công.

a. *Phương pháp giám sát nhật ký ứng dụng*

❖ Giám sát nhật ký truy cập (Access Log Monitoring)

Theo dõi các hoạt động truy cập vào ứng dụng, bao gồm thông tin như:

- Ai đã truy cập (danh tính người dùng).
- Thời gian truy cập (timestamp).
- Địa chỉ IP hoặc vị trí truy cập.

❖ Ứng dụng:

- Phát hiện các nỗ lực đăng nhập trái phép.
- Giám sát hành vi truy cập không thường xuyên vào các tệp hoặc tài nguyên nhạy cảm.

b. *Phân tích lỗi và ngoại lệ (Error and Exception Analysis)*

Thu thập và phân tích các lỗi và ngoại lệ được ghi lại trong nhật ký ứng dụng, từ đó phát hiện các hành vi bất thường hoặc sự cố bảo mật tiềm ẩn.

❖ Ứng dụng:

- Ghi lại các lỗi liên quan đến cơ sở dữ liệu (như truy vấn thất bại trong SQL).
- Theo dõi ngoại lệ trong mã ứng dụng, như lỗi không mong đợi liên quan đến xác thực.

c. *Lợi ích của App Log Monitoring*

- Bảo vệ ứng dụng khỏi các cuộc tấn công: SIEM giúp phát hiện các tấn công phổ biến như SQL Injection, Cross-Site Scripting (XSS), hoặc Directory Traversal thông qua phân tích các mẫu hành vi bất thường.
- Đảm bảo tính toàn vẹn và bảo mật của dữ liệu ứng dụng: Giảm thiểu nguy cơ lộ dữ liệu hoặc gián đoạn dịch vụ thông qua việc giám sát và phân tích nhật ký thời gian thực.

Object Access Auditing (Kiểm toán truy cập đối tượng)

Kiểm toán truy cập đối tượng là quy trình giám sát và ghi lại các hành vi truy cập vào các tài nguyên quan trọng trong hệ thống, như tệp tin, thư mục, cơ sở dữ liệu, hoặc thiết bị lưu trữ. Đây là yếu tố quan trọng trong việc đảm bảo rằng chỉ những người được ủy quyền mới có thể truy cập và tương tác với tài nguyên.

a. *Quy trình kiểm toán truy cập*

❖ Ghi lại lịch sử truy cập (Access History Logging)

Lưu lại thông tin chi tiết về:

- Ai đã truy cập vào tài nguyên (User Identity).
- Thời gian truy cập (Timestamp).

- Loại truy cập (Đọc, Ghi, Xóa).

❖ Ứng dụng:

- Ghi nhận ai đã thay đổi hoặc xóa tệp tin quan trọng.
- Theo dõi lịch sử truy cập vào tài liệu bí mật.

b. ***Phân tích hành vi truy cập (Access Behavior Analysis)***

Phân tích các mẫu hành vi truy cập để xác định các hoạt động bất thường.Ví dụ:

- Một nhân viên đột nhiên truy cập vào một lượng lớn tài liệu không liên quan đến vai trò công việc của họ.
- Một người dùng cố gắng truy cập vào thư mục được bảo vệ nhiều lần và thất bại.

c. ***Lợi ích của Object Access Auditing***

- Phát hiện các hành vi trái phép hoặc không hợp lệ: SIEM hỗ trợ phát hiện các hành vi xâm phạm bảo mật, từ đó ngăn chặn kịp thời.
- Tạo hồ sơ phục vụ điều tra và kiểm tra: Các bản ghi truy cập có thể được sử dụng làm bằng chứng trong các cuộc kiểm tra hoặc điều tra bảo mật.
- Cải thiện kiểm soát nội bộ: Hỗ trợ phát hiện và ngăn chặn các mối đe dọa từ bên trong, như nhân viên lạm dụng quyền hạn.

2.2.2 Chức năng của SIEM

Real-Time Monitoring (Theo dõi theo thời gian thực)

Real-Time Monitoring là một chức năng quan trọng trong hệ thống SIEM (Security Information and Event Management), cho phép theo dõi, giám sát các sự kiện bảo mật trong thời gian thực. Tính năng này được thiết kế nhằm phát hiện và phản ứng nhanh với các sự kiện đáng ngờ, giảm thiểu tác động của các mối đe dọa bảo mật lên hệ thống.

a. ***Tính năng của Real-Time Monitoring***

❖ Cảnh báo tự động

Hệ thống SIEM tự động phát hiện các sự kiện bất thường, dựa trên các quy tắc bảo mật (security rules) được thiết lập trước hoặc dựa trên học máy (machine learning), và gửi cảnh báo ngay lập tức đến quản trị viên.

Ví dụ: *Một tài khoản người dùng có hoạt động đăng nhập bất thường từ một địa điểm không quen thuộc sẽ kích hoạt cảnh báo.*

❖ Cách thức hoạt động:

- Phân tích thời gian thực: SIEM liên tục theo dõi dòng dữ liệu từ các nguồn nhật ký (logs) như tường lửa, thiết bị mạng, và máy chủ.
- Tích hợp hành động phản ứng: Sau khi phát hiện, hệ thống có thể tự động thực hiện các biện pháp bảo vệ như khóa tài khoản, cô lập thiết bị, hoặc chặn địa chỉ IP đáng ngờ.

❖ Lợi ích: Giúp quản trị viên phát hiện và phản ứng ngay lập tức với các mối đe dọa, giảm thiểu thiệt hại có thể xảy ra.

b. ***Cung cấp thông tin trực tiếp***

Thông qua bảng điều khiển (dashboard) tương tác, quản trị viên có thể theo dõi trạng thái bảo mật của toàn bộ hệ thống, bao gồm các sự kiện đáng ngờ, các điểm yếu tiềm tàng, và tình trạng hoạt động của thiết bị.

Tính năng hỗ trợ:

- Cập nhật trực tiếp các sự kiện đang diễn ra.
- Phân loại mức độ nghiêm trọng (low, medium, high) để ưu tiên xử lý.
- Hiển thị dữ liệu trực quan như biểu đồ, bản đồ mối đe dọa (threat maps), và các báo cáo phân tích.

c. ***Lợi ích của Real-Time Monitoring***

❖ Phát hiện và phản ứng ngay lập tức với mối đe dọa

Giảm thời gian phát hiện (Detection Time): Thay vì mất nhiều giờ hoặc ngày để phát hiện một cuộc tấn công, hệ thống SIEM có thể nhận diện mối đe dọa chỉ trong vài giây hoặc phút, nhờ khả năng xử lý dữ liệu theo thời gian thực.

Ví dụ: *Một cuộc tấn công brute-force sẽ được phát hiện khi có nhiều nỗ lực đăng nhập thất bại liên tục.*

❖ Hỗ trợ hành động nhanh chóng:

- Hệ thống có thể gửi cảnh báo qua email, tin nhắn SMS, hoặc các ứng dụng quản lý tập trung như Slack, Teams.
- Kết nối với các hệ thống bảo mật khác để thực hiện các biện pháp ngăn chặn như chặn truy cập hoặc cô lập máy chủ bị xâm nhập.

❖ Giảm thiểu thiệt hại tiềm tàng: Chặn đứng cuộc tấn công bằng việc phát hiện sớm giúp ngăn chặn các cuộc tấn công trước khi chúng gây ra hậu quả nghiêm trọng, chẳng hạn như đánh cắp dữ liệu hoặc làm gián đoạn hoạt động kinh doanh.

Ví dụ: *Một phần mềm mã độc (ransomware) có thể bị vô hiệu hóa trước khi mã hóa các tệp quan trọng.*

❖ Bảo vệ tài sản tổ chức:

- Ngăn chặn hành vi truy cập trái phép vào tài sản số quan trọng như cơ sở dữ liệu khách hàng hoặc thông tin tài chính.
- Đảm bảo hệ thống luôn sẵn sàng hoạt động, giảm thiểu thời gian chết (downtime).

❖ Cải thiện khả năng giám sát toàn diện: Toàn cảnh về bảo mật hệ thống: Với khả năng tập trung dữ liệu từ nhiều nguồn khác nhau, SIEM cung cấp cho quản trị viên cái nhìn toàn diện về trạng thái bảo mật, từ hoạt động của thiết bị đầu cuối đến lưu lượng mạng.

❖ Phân tích xu hướng và mối đe dọa: Thông qua việc ghi lại và phân tích các sự kiện, hệ thống có thể dự đoán các mối đe dọa tiềm ẩn trong tương lai.

❖ Nâng cao hiệu quả làm việc của đội ngũ an ninh

- Tự động hóa quy trình: Thay vì phải kiểm tra thủ công từng sự kiện, SIEM giúp đội ngũ bảo mật tập trung vào các sự cố quan trọng và giảm thiểu thời gian cho các công việc lặp đi lặp lại.
- Hỗ trợ ra quyết định nhanh:Bằng cách cung cấp thông tin chính xác và tức thì, hệ thống giúp quản trị viên đưa ra các hành động kịp thời.

User Activity Monitoring (Giám sát hoạt động người dùng)

Chức năng giám sát hoạt động người dùng của SIEM tập trung vào việc theo dõi và phân tích các hành vi của người dùng trong hệ thống. Điều này giúp phát hiện các hoạt động bất thường, từ các cuộc tấn công nội bộ đến các hành vi lạm dụng quyền hạn, qua đó bảo vệ tài sản tổ chức khỏi các rủi ro bảo mật.

a. Phương pháp giám sát

❖ Theo dõi hành động đăng nhập

- Ghi lại thời gian, địa điểm (IP, thiết bị), và phương thức đăng nhập của người dùng.
- Phát hiện các bất thường như đăng nhập từ vị trí địa lý bất thường hoặc sử dụng các tài khoản không hợp lệ.

❖ Giám sát truy cập dữ liệu

- Theo dõi dữ liệu mà người dùng truy cập, chỉnh sửa, tải xuống hoặc xóa khỏi hệ thống.
- Cảnh báo khi có hành động trái phép, ví dụ: nhân viên truy cập tệp dữ liệu không thuộc phạm vi công việc.

❖ Phân tích các thay đổi hệ thống

- Ghi nhận các thay đổi trong cấu hình hệ thống hoặc chính sách bảo mật do người dùng thực hiện.
- Cảnh báo khi các thay đổi này không được cấp phép hoặc vi phạm chính sách bảo mật.

b. Lợi ích

❖ Bảo vệ tổ chức khỏi các cuộc tấn công nội bộ: Phát hiện kịp thời các hành vi lạm dụng quyền hạn hoặc xâm nhập trái phép từ bên trong tổ chức.

Ví dụ: *Một nhân viên cố gắng trích xuất dữ liệu nhạy cảm để sử dụng cá nhân hoặc chuyển giao ra bên ngoài.*

❖ Hỗ trợ cải thiện chính sách bảo mật: Dựa trên các dữ liệu giám sát, tổ chức có thể xác định điểm yếu trong chính sách bảo mật hiện tại và thực hiện điều chỉnh phù hợp.

- ❖ Tăng cường khả năng kiểm toán: Cung cấp lịch sử chi tiết về hoạt động người dùng, hỗ trợ các cuộc điều tra hoặc kiểm toán tuân thủ.

Dashboards (Bảng điều khiển)

Bảng điều khiển trong SIEM cung cấp giao diện trực quan, giúp quản trị viên giám sát toàn diện các sự kiện bảo mật, chỉ số hiệu suất, và tình trạng của hệ thống. Đây là công cụ quan trọng để nắm bắt nhanh chóng các thông tin quan trọng.

a. Tính năng của bảng điều khiển

- ❖ Hiển thị thời gian thực: Cập nhật liên tục về trạng thái hệ thống, các sự kiện an ninh, và các cảnh báo quan trọng.

Ví dụ: *Quản trị viên có thể theo dõi trực tiếp lưu lượng mạng và phát hiện các hoạt động bất thường ngay lập tức.*

- ❖ Tùy chỉnh theo nhu cầu: Quản trị viên có thể tùy chỉnh giao diện để hiển thị các thông tin ưu tiên, chẳng hạn như các mối đe dọa tiềm tàng, sự cố đã được xử lý hoặc các xu hướng bảo mật dài hạn.
- ❖ Trực quan hóa dữ liệu: Cung cấp biểu đồ, đồ thị, và các bản đồ tương tác giúp dễ dàng phân tích và so sánh dữ liệu bảo mật.

b. Lợi ích

- ❖ Giúp theo dõi tình hình bảo mật dễ dàng hơn: Bảng điều khiển giúp đội ngũ bảo mật nắm bắt nhanh chóng trạng thái toàn bộ hệ thống, thay vì phải truy cập từng phần riêng lẻ.
- ❖ Cải thiện khả năng ra quyết định: Dựa trên dữ liệu hiển thị, quản trị viên có thể xác định các sự cố ưu tiên cần giải quyết và đưa ra phản ứng kịp thời.
- ❖ Hỗ trợ báo cáo nhanh: Dữ liệu từ bảng điều khiển có thể được xuất ra báo cáo để phục vụ cho các cuộc họp hoặc kiểm toán.

Reporting (Báo cáo)

Chức năng báo cáo của SIEM tạo ra các báo cáo chi tiết, cung cấp thông tin về tình hình bảo mật, hoạt động người dùng, và các sự kiện đáng chú ý. Báo cáo hỗ trợ tổ chức trong việc tuân thủ các tiêu chuẩn và quy định, cũng như cải thiện quản lý nội bộ.

a. Tính năng của chức năng báo cáo

- ❖ Tạo báo cáo tự động: SIEM có thể tự động tạo báo cáo định kỳ về các hoạt động bảo mật, giúp giảm bớt gánh nặng cho đội ngũ IT.

Ví dụ: *Báo cáo hàng tuần về số lượng cảnh báo bảo mật hoặc các mối đe dọa đã được xử lý.*

- ❖ Lưu trữ lịch sử sự kiện: Báo cáo cung cấp thông tin chi tiết về các sự kiện đã xảy ra, giúp tổ chức phân tích các mẫu hành vi hoặc điều tra các sự cố trước đó.

b. Lợi ích

- ❖ Hỗ trợ tuân thủ quy định: Cung cấp tài liệu cần thiết cho các cuộc kiểm tra tuân thủ như GDPR, ISO 27001, hoặc PCI DSS.
- ❖ Nâng cao nhận thức nội bộ: Báo cáo giúp đội ngũ quản lý và lãnh đạo hiểu rõ hơn về tình trạng bảo mật của tổ chức, từ đó đưa ra các quyết định chiến lược phù hợp.
- ❖ Hỗ trợ điều tra sự cố: Báo cáo lưu trữ dữ liệu chi tiết, tạo điều kiện thuận lợi cho việc điều tra pháp y hoặc xử lý sự cố sau này.

Threat Intelligence (Thông tin tình báo về mối đe dọa)

Chức năng này tích hợp các nguồn thông tin về mối đe dọa, giúp hệ thống SIEM phát hiện các mẫu tấn công mới và phản ứng hiệu quả hơn với các nguy cơ bảo mật.

a. Phương pháp tích hợp

- ❖ Thông tin từ các nguồn công cộng: Sử dụng dữ liệu từ các tổ chức bảo mật uy tín hoặc cộng đồng chia sẻ thông tin như Open Threat Exchange (OTX).
- ❖ Thông tin từ các nhà cung cấp thương mại: Sử dụng các dịch vụ cung cấp danh sách IP độc hại, mã độc, và các lỗ hổng mới từ các nhà cung cấp chuyên nghiệp.

b. Lợi ích

- ❖ Cải thiện khả năng phát hiện mối đe dọa mới: Với thông tin tình báo cập nhật, hệ thống SIEM có thể phát hiện và ngăn chặn các mối đe dọa mới mà không cần chờ sự cố xảy ra.
- ❖ Tăng cường khả năng phòng ngừa: Hệ thống có thể tự động cập nhật để bảo vệ tổ chức khỏi các mẫu tấn công phổ biến hoặc các nguy cơ bảo mật mới.

Incident Response (Phản ứng sự cố)

Phản ứng sự cố là một chức năng quan trọng của SIEM, giúp tổ chức xử lý hiệu quả các sự kiện bảo mật bằng cách cung cấp thông tin chi tiết và các hướng dẫn cần thiết.

a. Phương pháp phản ứng sự cố

- ❖ Quy trình phản ứng tiêu chuẩn: SIEM hỗ trợ xây dựng các quy trình phản ứng cho các tình huống cụ thể, như tấn công mạng, xâm nhập trái phép, hoặc mã độc.
- ❖ Hỗ trợ thông tin chi tiết: Cung cấp dữ liệu cụ thể như nguồn gốc tấn công, tài nguyên bị ảnh hưởng, và mức độ nghiêm trọng để đội ngũ an ninh có thể xử lý hiệu quả.

b. Lợi ích

- ❖ Giảm thiểu thời gian phản ứng: Hỗ trợ đội ngũ IT phát hiện và xử lý sự cố nhanh hơn, từ đó giảm thiểu thiệt hại và rủi ro kéo dài.
- ❖ Tăng cường khả năng phục hồi: Giúp tổ chức nhanh chóng khôi phục sau sự cố và duy trì hoạt động liên tục.
- ❖ Cải thiện quản lý sự cố: Dữ liệu từ quá trình phản ứng sự cố có thể được sử dụng để cải thiện quy trình, giảm nguy cơ tái diễn.

2.3 Lợi ích của việc xây dựng hệ thống giám sát

Hệ thống Quản lý thông tin và sự kiện bảo mật (SIEM - Security Information and Event Management) không chỉ cung cấp khả năng giám sát theo thời gian thực mà còn hỗ trợ ngăn chặn, phát hiện, và điều tra các sự cố bảo mật. Từ đó, SIEM đóng vai trò trung tâm trong việc bảo vệ hạ tầng CNTT của tổ chức trước các mối đe dọa ngày càng phức tạp.

Các lợi ích nổi bật của hệ thống SIEM:

- ❖ Tăng cường hiệu quả của đội ngũ an ninh và tối ưu hóa thời gian làm việc
    - SIEM tự động hóa quá trình phân tích sự kiện bảo mật, giảm thiểu khối lượng công việc thủ công của đội ngũ bảo mật.
    - Tích hợp các cảnh báo tự động, giúp nhóm bảo mật nhanh chóng tập trung vào các vấn đề quan trọng thay vì xử lý dữ liệu thô từ nhiều nguồn.
    - Hỗ trợ thông tin trực quan hóa qua bảng điều khiển (dashboards), tăng khả năng giám sát và phân tích trong thời gian ngắn.

Ví dụ thực tiễn:*Một tổ chức sử dụng SIEM có thể phát hiện 10 sự kiện đáng ngờ từ hàng triệu sự kiện nhật ký mỗi ngày, giúp đội ngũ bảo mật tập trung vào các mối đe dọa quan trọng nhất.*

❖ Ngăn chặn các mối đe dọa bảo mật tiềm ẩn trước khi chúng trở thành sự cố lớn

- SIEM sử dụng kỹ thuật tương quan sự kiện để xác định các mối đe dọa tiềm ẩn.
- Tích hợp các công cụ tình báo mối đe dọa (Threat Intelligence) giúp nhận diện các mẫu tấn công hoặc địa chỉ IP độc hại trước khi chúng gây thiệt hại.

Ví dụ thực tiễn:*Nếu SIEM phát hiện một địa chỉ IP liên tục thực hiện các hành động đăng nhập thất bại từ nhiều tài khoản, nó có thể tự động chặn địa chỉ IP đó và cảnh báo đội ngũ bảo mật.*

❖ Giảm chi phí bảo mật tổng thể cho tổ chức

- Tích hợp SIEM giúp tối ưu hóa quy trình bảo mật, giảm thiểu sự phụ thuộc vào nhiều công cụ riêng lẻ.
- Tự động hóa quy trình giảm bớt chi phí nhân lực và thời gian dành cho việc phân tích dữ liệu hoặc điều tra thủ công.
- Ngăn chặn sự cố từ sớm, giảm thiểu chi phí khắc phục hậu quả.

Lợi ích thực tế:*Theo nghiên cứu, tổ chức triển khai SIEM giảm được từ 30-40% chi phí dành cho việc quản lý và xử lý các sự cố an ninh so với hệ thống truyền thống.*

❖ Cung cấp hệ thống tốt hơn cho báo cáo, phân tích nhật ký và lưu trữ dữ liệu

- SIEM tạo báo cáo tự động, giúp đáp ứng yêu cầu kiểm toán và tuân thủ các tiêu chuẩn bảo mật như GDPR, PCI DSS, ISO 27001.
- Cung cấp kho lưu trữ dữ liệu tập trung, thuận tiện cho việc phân tích sự kiện và pháp y sau sự cố.

Ví dụ thực tiễn*:Trong một cuộc kiểm toán bảo mật, tổ chức có thể trích xuất dữ liệu từ SIEM để chứng minh các hoạt động giám sát đã được thực hiện liên tục, hỗ trợ việc tuân thủ quy định.*

❖ Giảm thiểu tác động của các cuộc tấn công bảo mật

- SIEM cung cấp cảnh báo theo thời gian thực, giúp giảm thời gian phát hiện và phản ứng với các cuộc tấn công.
- Với khả năng ghi lại chi tiết hành vi của kẻ tấn công, hệ thống giúp giảm thiệt hại và tăng tốc độ khắc phục.

Ví dụ thực tiễn:*Một tổ chức phát hiện mã độc ransomware trong vòng 5 phút nhờ SIEM, từ đó nhanh chóng cô lập hệ thống bị tấn công trước khi mã độc lây lan.*

2.4 Các hệ thống giám sát mạng phổ biến hiện nay

2.4.1 SPLUNK

Splunk là một trong những giải pháp SIEM (Security Information and Event Management) mạnh mẽ nhất hiện nay, được phát triển bởi Splunk Inc**.** và phát hành lần đầu vào năm 2003. Hệ thống này được thiết kế để xử lý và phân tích dữ liệu lớn từ nhiều nguồn khác nhau, từ đó cung cấp thông tin chi tiết về an ninh mạng và hoạt động của hệ thống.

Hình 2.3:Hệ thống Splunk

a. Khả năng nổi bật của Splunk

❖ Xử lý và phân tích dữ liệu lớn

Splunk được thiết kế để xử lý các luồng dữ liệu lớn một cách hiệu quả, từ thu thập đến phân tích và lưu trữ. Các tính năng chính bao gồm:

- Thu thập và xử lý đa dạng nguồn dữ liệu: Splunk có khả năng thu thập dữ liệu từ nhiều loại nguồn như nhật ký hệ thống, sự kiện mạng, hoạt động ứng dụng, cơ sở dữ liệu, thiết bị IoT, và thậm chí cả dịch vụ đám mây.
- Công nghệ xử lý phân tán: Splunk sử dụng kiến trúc phân tán với khả năng xử lý dữ liệu lớn trên nhiều máy chủ, giúp tăng hiệu suất xử lý và giảm thời gian phản hồi.
- Tìm kiếm nhanh: Sử dụng Splunk Search Processing Language (SPL), người dùng có thể truy vấn và phân tích dữ liệu hàng triệu sự kiện trong vài giây.
- Phân tích dữ liệu tiên tiến: Splunk hỗ trợ các công cụ học máy và phân tích dữ liệu để tự động xác định các mẫu, xu hướng, và phát hiện bất thường.

Ví dụ thực tiễn:*Trong một hệ thống giám sát giao thông thông minh, Splunk có thể phân tích dữ liệu từ hàng ngàn cảm biến và camera theo thời gian thực để phát hiện các khu vực ùn tắc hoặc tai nạn.*

❖ Tích hợp linh hoạt với nhiều nguồn dữ liệu

Splunk là một hệ thống SIEM đa năng với khả năng tích hợp gần như mọi loại dữ liệu nhờ khả năng hỗ trợ:

- Tích hợp trực tiếp: Hỗ trợ nhiều giao thức và định dạng dữ liệu phổ biến như syslog, SNMP, JSON, XML, CSV, và giao thức REST API.
- Tích hợp với hệ sinh thái DevOps: Splunk dễ dàng kết hợp với các công cụ CI/CD như Jenkins, Kubernetes, hoặc Docker, giúp theo dõi toàn bộ quy trình phát triển và vận hành ứng dụng.
- Hỗ trợ dịch vụ đám mây: Tích hợp với AWS, Microsoft Azure, và Google Cloud Platform để thu thập và phân tích dữ liệu từ các ứng dụng hoặc cơ sở hạ tầng đám mây.

Ví dụ thực tiễn:*Một tổ chức tài chính tích hợp Splunk với dịch vụ đám mây AWS và các thiết bị bảo mật Cisco để giám sát hoạt động mạng, phát hiện giao dịch đáng ngờ và tuân thủ các tiêu chuẩn bảo mật như PCI-DSS.*

❖ Giao diện người dùng mạnh mẽ và trực quan

Splunk cung cấp giao diện đồ họa dễ sử dụng, giúp quản trị viên và các nhóm an ninh nhanh chóng hiểu rõ tình hình hệ thống. Một số tính năng nổi bật:

- Dashboard trực quan: Các bảng điều khiển được thiết kế tùy chỉnh, hiển thị thông tin thời gian thực với các biểu đồ tương tác như heatmap, timeline, hoặc pie chart.
- Quản lý cảnh báo: Splunk cho phép tạo và quản lý cảnh báo dễ dàng dựa trên các điều kiện được định nghĩa trước (ví dụ: số lượng đăng nhập thất bại liên tiếp).
- Công cụ SPL mạnh mẽ: SPL giúp người dùng truy vấn dữ liệu với độ chính xác cao mà không cần hiểu biết sâu về lập trình.

Ví dụ: *Quản trị viên sử dụng bảng điều khiển của Splunk để theo dõi đăng nhập không hợp lệ trong thời gian thực. Khi phát hiện có một IP thực hiện hàng trăm lần đăng nhập thất bại, hệ thống sẽ gửi cảnh báo ngay lập tức qua email.*

❖ Khả năng mở rộng vượt trội

Splunk được xây dựng với kiến trúc phân tán, cho phép mở rộng theo chiều ngang một cách dễ dàng.

- Mở rộng quy mô**:** Người dùng có thể tăng số lượng máy chủ hoặc chuyển sang Splunk Cloud để xử lý lượng lớn dữ liệu mà không ảnh hưởng đến hiệu suất.
- Hỗ trợ nhiều mô hình triển khai: Splunk có thể được triển khai tại chỗ (on-premises), trên đám mây, hoặc trong môi trường lai (hybrid), đáp ứng nhu cầu linh hoạt của tổ chức.

Ví dụ thực tiễn:*Một công ty thương mại điện tử lớn mở rộng hệ thống Splunk để giám sát hàng triệu giao dịch mỗi ngày trong mùa mua sắm.*

b. Ưu điểm của Splunk

❖ Dễ sử dụng:

- Giao diện thân thiện giúp người dùng không cần kỹ năng lập trình vẫn có thể sử dụng thành thạo.

- Tích hợp tự động hóa cho nhiều tác vụ, từ thu thập dữ liệu đến tạo báo cáo.

❖ Tích hợp đa dạng: Splunk hỗ trợ hơn 1.000 ứng dụng và plugin thông qua Splunkbase, bao gồm các công cụ bảo mật, giám sát hệ thống, và phân tích hiệu năng.

❖ Phân tích mạnh mẽ: Tích hợp các thuật toán học máy, hỗ trợ phát hiện và dự đoán các mối đe dọa tiềm ẩn dựa trên hành vi bất thường.

❖ Tùy chỉnh linh hoạt: Splunk cho phép tùy chỉnh báo cáo, cảnh báo, và các quy trình phân tích theo nhu cầu cụ thể.

❖ Cộng đồng hỗ trợ : Splunk có một cộng đồng phát triển mạnh mẽ với nhiều tài liệu, hướng dẫn, và các ứng dụng miễn phí được chia sẻ trên Splunkbase.

c. Nhược điểm của Splunk

❖ Chi phí cao:

Splunk là một trong những hệ thống SIEM đắt đỏ nhất hiện nay. Chi phí bao gồm:

- Bản quyền phần mềm (theo dung lượng dữ liệu và thời gian lưu trữ).
- Tài nguyên phần cứng (CPU, RAM, lưu trữ).
- Chi phí nhân lực vận hành.

Ví dụ:*Một doanh nghiệp lớn cần xử lý 1PB dữ liệu mỗi ngày với Splunk có thể tiêu tốn hàng triệu USD/năm cho giấy phép và cơ sở hạ tầng.*

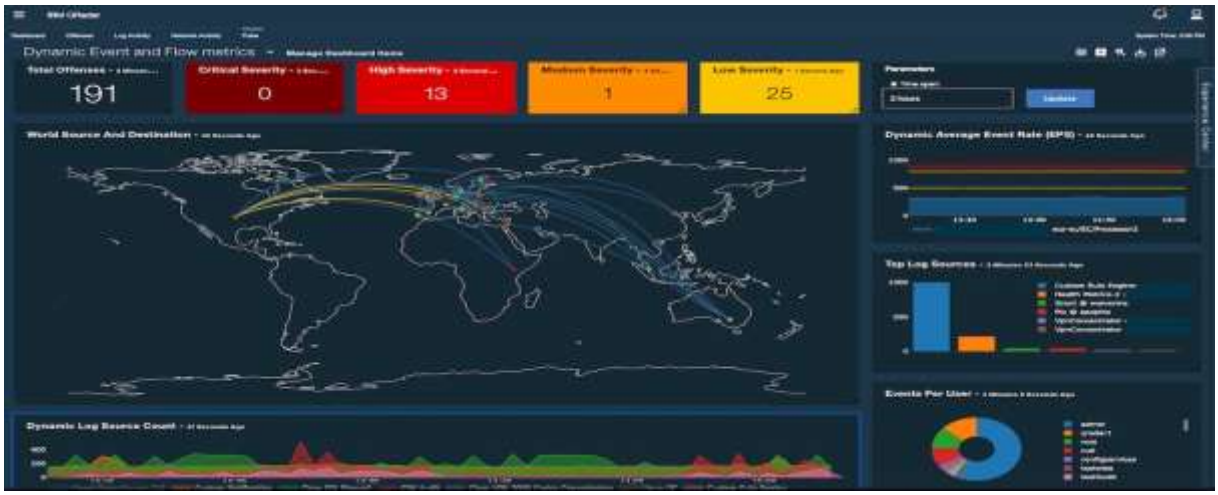
❖ Đòi hỏi tài nguyên lớn: Splunk yêu cầu tài nguyên phần cứng mạnh để đảm bảo hiệu suất khi xử lý dữ liệu lớn. Điều này có thể là một rào cản đối với các tổ chức nhỏ.

❖ Phức tạp khi triển khai:

- Việc triển khai Splunk yêu cầu đội ngũ kỹ thuật có kinh nghiệm cao, đặc biệt khi tùy chỉnh hoặc tích hợp hệ thống phức tạp.
- Một số tính năng nâng cao như học máy đòi hỏi kỹ năng chuyên môn sâu.

2.4.2 IBM QRadar

IBM QRadar là một hệ thống SIEM được phát triển ban đầu bởi Q1 Labs vào năm 2005 và sau đó được IBM mua lại vào năm 2011. QRadar là một trong những giải pháp SIEM nổi bật trong ngành bảo mật, chuyên phát hiện các mối đe dọa an ninh mạng và hỗ trợ quản lý sự cố tự động.



Hình 2.4:Hệ thống IBM QRadar

a. Khả năng nổi bật của IBM QRadar

❖ Thu thập và phân tích dữ liệu từ nhiều nguồn khác nhau:

IBM QRadar có khả năng thu thập dữ liệu từ đa dạng nguồn, bao gồm các thiết bị mạng (firewall, router), ứng dụng, hệ thống bảo mật, dịch vụ đám mây và cơ sở hạ tầng trên toàn cầu.

- Khả năng thu thập dữ liệu linh hoạt: QRadar có thể thu thập và tổng hợp dữ liệu từ nhiều dạng như nhật ký hệ thống, sự kiện bảo mật, và lưu trữ từ các hệ thống mạng, giúp cung cấp cái nhìn toàn diện về tình hình bảo mật.
- Khả năng phân tích mạnh mẽ: QRadar có thể phân tích dữ liệu từ các nguồn khác nhau và đánh giá nguy cơ mối đe dọa một cách tự động, hỗ trợ các tổ chức trong việc phát hiện các sự kiện bảo mật quan trọng.

Ví dụ thực tiễn:*Một tổ chức tài chính sử dụng IBM QRadar để thu thập và phân tích dữ liệu từ firewall, IDS/IPS, hệ thống giám sát giao dịch và các ứng dụng tài chính. QRadar nhanh chóng phát hiện các mẫu tấn công tinh vi hoặc giao dịch bất thường, cung cấp thông tin chi tiết để đội ngũ bảo mật phản ứng kịp thời.*

❖ Tích hợp AI và khả năng tự động hoá bảo mật

Một trong những tính năng nổi bật của IBM QRadar là tích hợp trí tuệ nhân tạo (AI) và các thuật toán học máy để tăng cường khả năng phát hiện và phân tích mối đe dọa:

- Phát hiện mối đe dọa thông minh: AI và học máy giúp QRadar phân tích dữ liệu và xác định các mẫu bất thường, giúp giảm thiểu thời gian phát hiện các cuộc tấn công.
- Tự động hoá phản ứng bảo mật: QRadar có thể tự động kích hoạt các hành động phản ứng khi phát hiện mối đe dọa, chẳng hạn như tạo cảnh báo, cập nhật cấu hình hệ thống, hoặc cách ly các thiết bị bị ảnh hưởng.
- Tích hợp với các công cụ bảo mật khác: QRadar tích hợp với các công cụ bảo mật khác để cung cấp khả năng giám sát và phản ứng linh hoạt và hiệu quả trong toàn bộ hệ thống bảo mật.

Ví dụ thực tiễn:*Khi phát hiện hành vi lạ từ một địa chỉ IP trong mạng, QRadar có thể tự động phân tích thông tin từ nhiều nguồn dữ liệu và đưa ra quyết định để ngừng hoặc giới hạn quyền truy cập từ địa chỉ đó, đồng thời tạo báo cáo chi tiết cho quản trị viên.*

❖ Phân tích mối đe dọa nhanh chóng

IBM QRadar cung cấp khả năng phân tích nhanh chóng các mối đe dọa bảo mật thông qua các công cụ phân tích nâng cao và khả năng xử lý dữ liệu mạnh mẽ:

- Phân tích các sự kiện an ninh trong thời gian thực: QRadar cung cấp khả năng phân tích các sự kiện bảo mật theo thời gian thực, giúp phát hiện và phản ứng nhanh chóng trước các mối đe dọa.
- Tìm kiếm và phân tích sự kiện một cách nhanh chóng: QRadar cho phép quản trị viên truy vấn và phân tích các sự kiện bảo mật trong thời gian thực, giúp phát hiện các hành vi đáng ngờ hoặc các cuộc tấn công tiềm ẩn.

Ví dụ thực tiễn:*QRadar có thể nhanh chóng phát hiện một cuộc tấn công DDoS dựa trên phân tích lưu lượng mạng và các cảnh báo từ hệ thống tường lửa. Thông qua khả năng phân tích nhanh chóng, QRadar giúp đội ngũ bảo mật đưa ra các biện pháp đối phó hiệu quả.*

b. Ưu điểm của IBM QRadar

- ❖ Tích hợp AI và học máy: IBM QRadar tích hợp AI và học máy để phân tích và phát hiện mối đe dọa một cách nhanh chóng và chính xác. Điều này giúp giảm thiểu thời gian phát hiện và phản ứng với các cuộc tấn công.
- ❖ Khả năng tự động hoá cao: QRadar cung cấp khả năng tự động hoá trong việc phát hiện và phản ứng với các mối đe dọa, giúp tăng cường hiệu quả bảo mật và giảm tải cho các đội ngũ an ninh.
- ❖ Phân tích mối đe dọa nhanh chóng: QRadar có khả năng phân tích dữ liệu từ nhiều nguồn và cung cấp thông tin chi tiết về các mối đe dọa trong thời gian thực, giúp tổ chức phản ứng kịp thời.
- ❖ Tích hợp mạnh mẽ với các công cụ bảo mật: QRadar có khả năng tích hợp với nhiều công cụ bảo mật khác như IDS/IPS, hệ thống tường lửa, và các công cụ phân tích an ninh khác, giúp xây dựng một hệ sinh thái bảo mật mạnh mẽ và toàn diện.
- ❖ Cung cấp báo cáo chi tiết và khả năng phân tích nâng cao: QRadar cung cấp các báo cáo chi tiết và phân tích các sự kiện bảo mật theo các tiêu chí khác nhau, giúp các tổ chức dễ dàng tuân thủ các yêu cầu bảo mật và phân tích tình trạng an ninh hệ thống.

c. Nhược điểm của IBM QRadar

- ❖ Yêu cầu kiến thức chuyên môn cao: QRadar là một hệ thống SIEM phức tạp, và việc triển khai cũng như cấu hình đòi hỏi đội ngũ kỹ thuật có kiến thức chuyên môn sâu. Các tính năng nâng cao như cấu hình AI và tự động hoá yêu cầu kinh nghiệm và sự hiểu biết về công cụ.

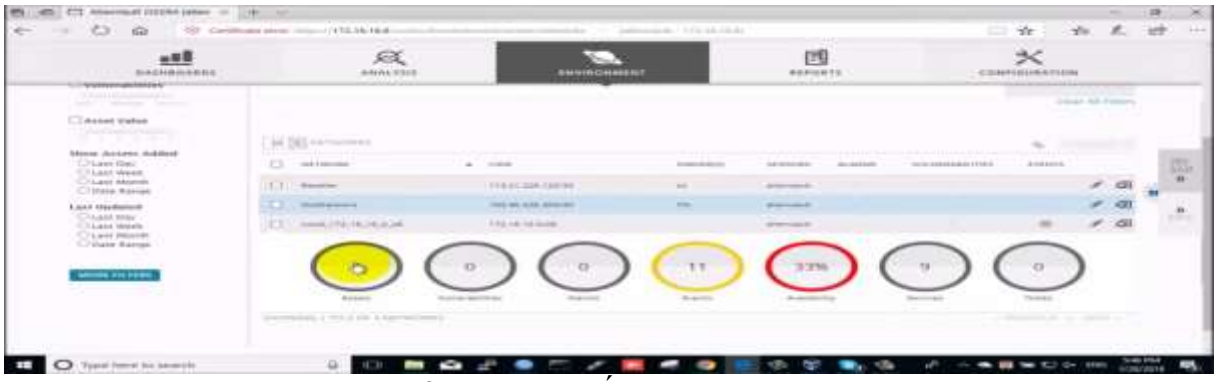
Ví dụ:*Việc cấu hình các cảnh báo và quy trình phản ứng tự động yêu cầu phải hiểu rõ về các mô hình dữ liệu và phân tích hành vi trong QRadar, điều này có thể là một thách thức đối với các tổ chức thiếu chuyên môn.*

- ❖ Chi phí triển khai và bảo trì cao: QRadar có chi phí khá cao, đặc biệt khi triển khai trên quy mô lớn. Các chi phí liên quan đến bản quyền phần mềm, tài nguyên phần cứng, và nhân lực vận hành có thể tăng nhanh.

Ví dụ:*Một tổ chức phải chi trả hàng trăm ngàn đến triệu đô la để triển khai và duy trì QRadar cho hệ thống với hàng triệu sự kiện bảo mật mỗi ngày.*

2.4.3 AlienVault OSSIM

AlienVault OSSIM (*Open Source Security Information and Event Management*) được phát hành vào năm 2003 bởi AT&T Cybersecurity (trước đây là AlienVault). OSSIM là một giải pháp SIEM mã nguồn mở, giúp các tổ chức quản lý logs, phát hiện mối đe dọa, quản lý sự cố bảo mật, và tích hợp với các công cụ bảo mật khác. OSSIM cung cấp các tính năng cơ bản của SIEM, nhưng với một cách tiếp cận đơn giản và dễ sử dụng, phù hợp với các tổ chức không cần các hệ thống SIEM phức tạp.



Hình 2.5 :Hệ thống AlienVault OSSIM

Khả năng nổi bật của AlienVault OSSIM

a. ***Quản lý logs và phân tích dữ liệu bảo mật***

AlienVault OSSIM cung cấp khả năng thu thập và phân tích logs từ nhiều nguồn khác nhau như thiết bị mạng (firewall, IDS/IPS), máy chủ, ứng dụng và các dịch vụ đám mây. OSSIM tổng hợp các sự kiện bảo mật và cung cấp thông tin chi tiết về các mối đe dọa, giúp phát hiện các hoạt động đáng ngờ hoặc các cuộc tấn công.

- Khả năng thu thập logs mạnh mẽ: OSSIM hỗ trợ nhiều giao thức và định dạng khác nhau như syslog, SNMP, và các dạng log phổ biến khác, giúp thu thập dữ liệu từ các thiết bị và hệ thống mạng đa dạng.
- Phân tích sự kiện bảo mật: OSSIM có thể phân tích và phát hiện các mối đe dọa bảo mật thông qua các quy tắc (rules) và các thuật toán phân tích. Các mối đe dọa được xác định và báo cáo để người quản trị có thể phản ứng kịp thời.

Ví dụ thực tiễn:*Một công ty nhỏ sử dụng AlienVault OSSIM để thu thập logs từ firewall, các máy chủ web, và các ứng dụng phần mềm. OSSIM tự động phân tích và phát hiện các hành vi đáng ngờ, như các lần đăng nhập trái phép hoặc sự thay đổi cấu hình không mong muốn.*

b. Tích hợp thông tin tình báo bảo mật (Threat Intelligence)

AlienVault OSSIM tích hợp thông tin tình báo bảo mật từ các nguồn khác nhau, giúp cải thiện khả năng phát hiện các mối đe dọa. OSSIM hỗ trợ các nguồn thông tin tình báo công cộng và thương mại, giúp hệ thống phát hiện các mối đe dọa chưa biết và các cuộc tấn công mới.

- ❖ Tích hợp dễ dàng: OSSIM có thể dễ dàng tích hợp với các dịch vụ tình báo mối đe dọa như AlienVault Open Threat Exchange (OTX), một cộng đồng chia sẻ thông tin về các mối đe dọa.
- ❖ Cập nhật thường xuyên: Các thông tin tình báo được cập nhật liên tục, giúp OSSIM nhận diện các mẫu tấn công mới và các mối đe dọa chưa được xác định.

Ví dụ thực tiễn:*OSSIM có thể nhận thông tin từ OTX về các địa chỉ IP độc hại và tự động kiểm tra xem các địa chỉ đó có xuất hiện trong các logs của hệ thống không. Nếu có, OSSIM sẽ kích hoạt cảnh báo và thông báo cho người quản trị.*

c. Quản lý sự cố bảo mật

AlienVault OSSIM cung cấp các công cụ giúp quản lý sự cố bảo mật, bao gồm khả năng tạo các cảnh báo, theo dõi sự kiện bảo mật, và cung cấp các công cụ phản ứng tự động. Điều này giúp các đội ngũ bảo mật phản ứng nhanh chóng và chính xác khi có sự cố xảy ra.

- ❖ Quản lý cảnh báo: OSSIM có thể tạo cảnh báo tự động khi phát hiện các sự kiện bảo mật quan trọng, giúp người quản trị có thể nhanh chóng xử lý các sự cố.
- ❖ Phản ứng tự động: OSSIM hỗ trợ các quy trình phản ứng tự động, giúp giảm thời gian phản ứng và giảm thiểu thiệt hại do các mối đe dọa.

Ví dụ thực tiễn:*Khi OSSIM phát hiện một cuộc tấn công DDoS, nó có thể tạo cảnh báo và tự động thực hiện các biện pháp như chặn lưu lượng từ các địa chỉ IP bị tấn công hoặc cô lập các máy chủ bị ảnh hưởng.*

Ưu điểm và nhược điểm của hệ thống

a. Ưu điểm của AlienVault OSSIM

- ❖ Mã nguồn mở và dễ tiếp cận: AlienVault OSSIM là một hệ thống SIEM mã nguồn mở, giúp giảm chi phí triển khai và bảo trì. Nó dễ tiếp cận và có thể được tùy

chỉnh theo nhu cầu của tổ chức, giúp các doanh nghiệp vừa và nhỏ dễ dàng triển khai mà không cần chi phí quá cao.

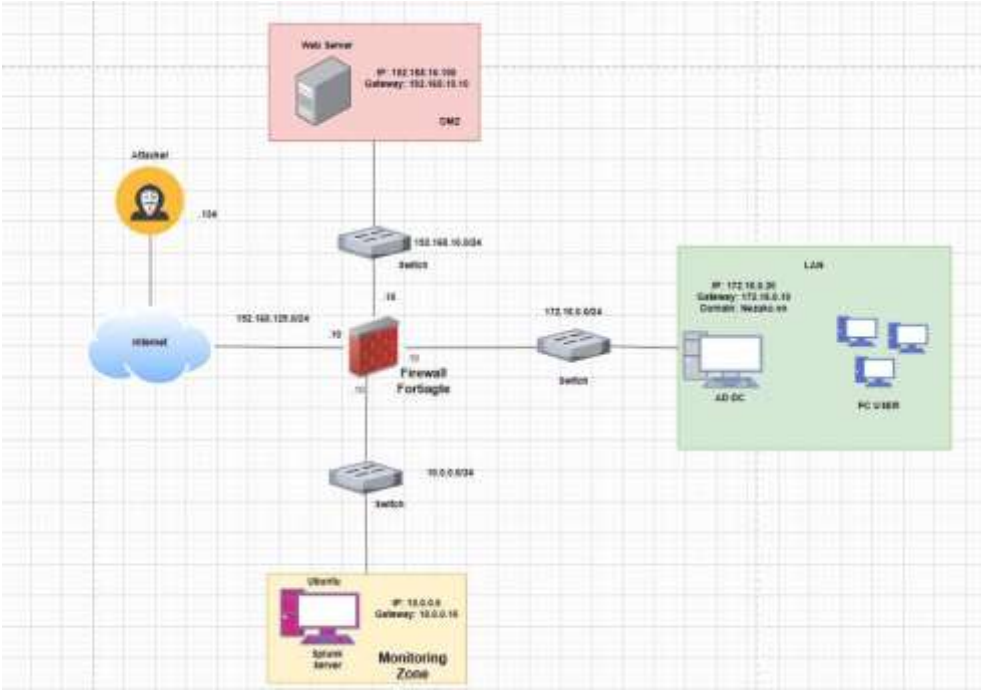
- Tích hợp thông tin tình báo bảo mật: OSSIM hỗ trợ tích hợp với các nguồn thông tin tình báo bảo mật nổi tiếng, bao gồm AlienVault Open Threat Exchange (OTX), giúp cung cấp khả năng phát hiện mối đe dọa tốt hơn thông qua các dữ liệu từ cộng đồng bảo mật.
- Dễ sử dụng và triển khai: AlienVault OSSIM dễ sử dụng, với giao diện người dùng đơn giản và dễ tiếp cận. Nó không yêu cầu kiến thức chuyên môn sâu để triển khai và vận hành, điều này làm cho nó phù hợp cho các tổ chức vừa và nhỏ hoặc các tổ chức không cần hệ thống SIEM phức tạp.
- Tích hợp với nhiều công cụ bảo mật khác: OSSIM hỗ trợ tích hợp với các công cụ bảo mật khác nhau như IDS/IPS, firewall, và các ứng dụng bảo mật khác, giúp tạo thành một hệ sinh thái bảo mật hoàn chỉnh và linh hoạt.

b. Nhược điểm của AlienVault OSSIM

- Khả năng mở rộng hạn chế: Mặc dù AlienVault OSSIM rất mạnh mẽ trong môi trường nhỏ và vừa, nhưng khả năng mở rộng của nó có thể bị hạn chế khi xử lý dữ liệu lớn hoặc triển khai trên các tổ chức quy mô lớn. Các tính năng nâng cao và phân tích phức tạp hơn có thể cần các công cụ bổ sung.
- Tính năng hạn chế so với các hệ thống SIEM cao cấp: AlienVault OSSIM cung cấp các tính năng cơ bản của SIEM, nhưng không có các công cụ phân tích và học máy tiên tiến như các giải pháp SIEM cao cấp khác (ví dụ: Splunk, QRadar). Điều này có thể làm giảm khả năng phát hiện các mối đe dọa tinh vi hoặc mối đe dọa chưa biết.
- Yêu cầu cấu hình và quản lý đúng cách: Dù là mã nguồn mở và dễ sử dụng, OSSIM vẫn yêu cầu người dùng có kiến thức cơ bản về bảo mật và hệ thống để cấu hình đúng cách. Việc không cấu hình đúng có thể dẫn đến các cảnh báo sai hoặc thiếu sót trong việc phát hiện mối đe dọa.

CHƯƠNG 3 : TRIỂN KHAI THỰC NGHIỆM HỆ THỐNG PHÁT HIỆN TẤN CÔNG

3.1 Mô hình thực nghiệm đề xuất



Hình 3.1: Mô Hình Đề Xuất Thực Nghiệm

Bảng 3:Bảng mô tả thiết bị

Tên thiết bị	IP	SUBNET MASK	GATEWAY
Web Server(Windowserver 2019)	VMNET 3: 192.168.10.100	255.255.255.0	192.168.100.10
Attacker (Kali)	Nat: 192.168.129.134	255.255.255.0	
Firewall (Fortigate)	Nat :192.168.129.10 VMNET 2: 10.0.0.10 VMNET 3: 192.168.10.10 VMNET 4: 172.16.0.10	255.255.255.0	
Splunk Server (Ubuntu)	VMNET 2: 10.0.0.5	255.255.255.0	10.0.0.10
Domain Controller (Windowserver 2019)	VMNET 4: 172.16.0.20	255.255.255.0	172.16.0.10

3.2 Công cụ đề xuất với mô hình thực nghiệm

Splunk là một giải pháp SIEM (Security Information and Event Management) tiên tiến, được phát triển bởi Splunk Inc., lần đầu ra mắt vào năm 2003. Công cụ này nổi bật với khả năng xử lý và phân tích dữ liệu lớn từ nhiều nguồn khác nhau, giúp cung cấp thông tin chi tiết về an ninh mạng cũng như hiệu suất hoạt động của hệ thống. Với tính năng mạnh mẽ và linh hoạt, Splunk là lựa chọn lý tưởng cho mô hình thực nghiệm trong đồ án lần này của chúng em.

3.2.1 Thành phần chính

Hình 3.2: Thành phần chính của SPLUNK

❖ **Forwarder**

Forwarder chính là một Agent Tool, một ứng dụng phần mềm mà ta sẽ cài trên các thiết bị muốn Splunk giám sát, có thể là các hosts, các Network Devices, Server … Khi cài thành công, lúc này các Logs sẽ được tự động gửi đến Splunk. Đối với hệ điều hành máy tính, Splunk có thể Collecting Logs từ Hệ điều hành Windows và Linux. Để biết trên Windows hay Linux, Splunk thu thập những Logs gì, thì đối với logs của Windows sẽ là Windows Event Logs, System Logs. Đối với Linux sẽ là Kernel, HTTPD, Authentication Logs.

❖ **Indexer**

Indexer chức năng chủ yếu là giúp Splunk lưu trữ Logs thu thập từ các thiết bị để chuẩn hóa và xử lý. Ta có một vấn đề là các file Logs thường có rất nhiều định dạng, có file đã chuẩn hóa, file chưa chuẩn hóa . Indexer sẽ giúp giải quyết cái vấn đề này bằng cách sẽ chuẩn hóa các file Logs thành một định dạng chung được định nghĩa là Field – Value.

Ví dụ: *IP – 192.168.1.100, Port – 80, Host – PC1, … Sau khi đã chuẩn hóa, Splunk có thể xử lý các Logs đó một cách hiệu quả.*

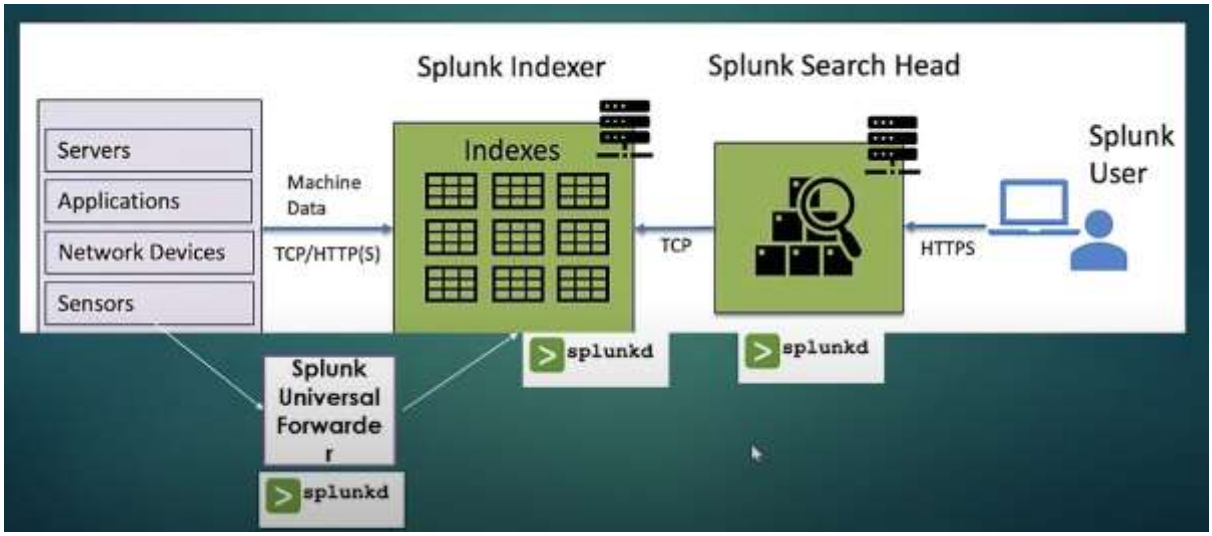
Một thứ đặc biệt của Splunk đó chính là kho lưu trữ được phân chia hợp lý mặc dù không cần sao lưu tất cả dữ liệu Splunk, vì phần lớn dữ liệu có thể có giá trị thấp. Thế nên Splunk cung cấp một hệ thống để chuyển đổi dữ liệu giữa bốn buckets lưu trữ, đại diện cho các giai đoạn khác nhau trong vòng đời dữ liệu:

- Hot Bucket: Chứa dữ liệu mới nhất mà Splunk thu thập. Splunk có thể truy vấn dữ liệu để phân tích tại đây.
- Warm Bucket:Lưu trữ dữ liệu cũ được chuyển sang từ Hot Bucket khi khi đạt ngưỡng dung lượng hoặc thời gian cấu hình.Khi đạt ngưỡng tiếp theo , dữ liệu sẽ tiếp tục được đẩy sang Cold Bucket.Cả Hot và Warm Bucket cần bộ nhớ mạnh, tốc độ cao để đảm bảo truy vấn nhanh.
- Cold Bucket: Tại đây sẽ là nơi lưu từ data cũ của Warm đẩy qua. Dữ liệu có thể bị xóa hoặc lưu trữ lâu dài vào Frozen Bucket, tùy theo nhu cầu doanh nghiệp.
- Frozen Bucket: Dữ liệu được lưu trữ lâu dài, không thể truy vấn trực tiếp.Muốn truy cập phải sử dụng script do Splunk cung cấp và khôi phục dữ liệu để tìm kiếm.

❖ **Search Head**

Search Head: Phần Search Head là một phần rất quan trọng trong hệ thống sẽ giúp cho ta có thể truy vấn và tìm kiếm dữ liệu trong các File Logs đã thu thập được thông qua một ngôn ngữ truy vấn được gọi tắt là SPL (hay còn gọi là Splunk Process Language).

3.2.2 Quy trình hoạt động



Hình 3.3 : Quy trình hoạt động của Splunk

Dữ liệu máy (machine data) từ các nguồn như máy chủ, ứng dụng, thiết bị mạng, hoặc cảm biến được Universal Forwarder thu thập và gửi đến Indexer qua giao thức TCP/HTTP(S). Indexer xử lý và lập chỉ mục dữ liệu để lưu trữ, cho phép dễ dàng tìm kiếm và phân tích. Splunk Search Head đóng vai trò giao diện người dùng, nơi các truy vấn được thực hiện và kết quả phân tích được hiển thị qua giao diện HTTPS. Người dùng cuối có thể tạo báo cáo, biểu đồ, và dashboard dựa trên dữ liệu được lập chỉ mục, hỗ trợ giám sát và phân tích dữ liệu hiệu quả

3.2.3 Một số thuật toán học máy hiện nay

Thuật toán học máy (Machine Learning Algorithm) là một phương pháp tính toán trong trí tuệ nhân tạo, cho phép máy tính học hỏi từ dữ liệu mà không cần phải lập trình chi tiết các quy tắc cụ thể. Mục tiêu của thuật toán học máy là phát hiện các mẫu trong dữ liệu, từ đó tạo ra các mô hình dự đoán hoặc phân loại có thể áp dụng cho các dữ liệu chưa thấy.

❖ Học máy thường có hai loại:

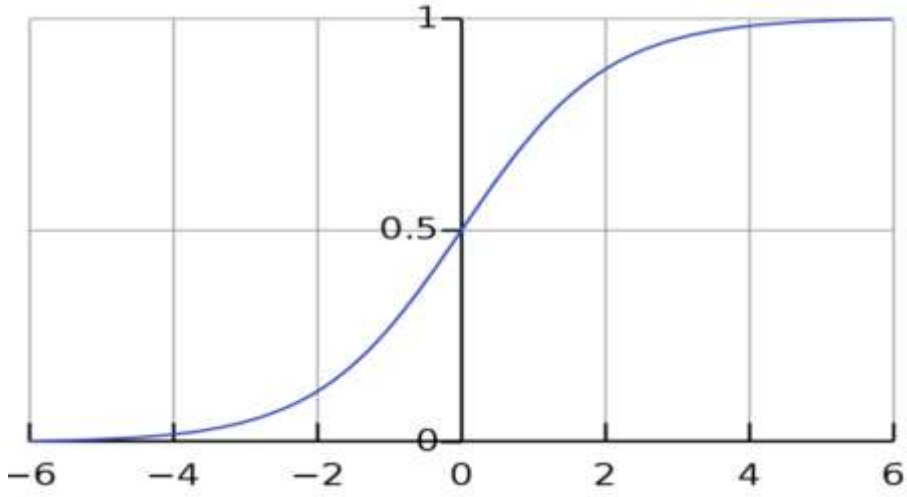
- Học giám sát (Supervised Learning): Dữ liệu huấn luyện đã được gắn nhãn (label) sẵn. Thuật toán học máy sử dụng các dữ liệu này để học và dự đoán các nhãn cho dữ liệu mới. Ví dụ: hồi quy tuyến tính, hồi quy logistic, cây quyết định, máy vector hỗ trợ (SVM), v.v.

- Học không giám sát (Unsupervised Learning): Dữ liệu huấn luyện không có nhãn, và thuật toán học máy phải tìm kiếm các mẫu hoặc cấu trúc trong dữ liệu mà không có sự hướng dẫn. Ví dụ: phân cụm (clustering), giảm chiều (dimensionality reduction), v.v.

a. Thuật toán hồi quy logistic (Logistic Regression)

Hồi quy logistic là một thuật toán học máy được sử dụng chủ yếu trong các bài toán phân loại, đặc biệt là phân loại nhị phân. Mục tiêu của hồi quy logistic là dự đoán xác suất của một sự kiện xảy ra, và kết quả của mô hình sẽ nằm trong khoảng từ 0 đến 1. Điều này rất hữu ích khi chúng ta cần phân loại các đối tượng thành hai nhóm, ví dụ như xác định một email là spam hay không spam, dự đoán bệnh nhân có nguy cơ mắc bệnh hay không, hoặc phân loại khách hàng tiềm năng.

Khác với hồi quy tuyến tính, trong hồi quy logistic, giá trị đầu ra được tính toán thông qua hàm logistic (còn gọi là hàm sigmoid) , giúp chuyển đổi giá trị thực tế thành xác suất**.** Hàm Sigmoid nhận đầu vào là một giá trị z bất kỳ, và trả về đầu ra là một giá trị xác suất nằm trong khoảng [0,1].



Hình 3.4: Biểu đồ biểu diễn thuật toán

Hàm sigmoid được sử dụng trong thuật toán hồi quy logistic để ánh xạ giá trị tuyến tính của mô hình thành một xác suất trong khoảng từ 0 đến 1. Trong Splunk Machine Learning Toolkit (MLTK), khi áp dụng cho phát hiện tấn công DoS, hàm sigmoid đóng vai trò chính trong việc phân loại sự kiện mạng dựa trên dữ liệu đầu vào.

Dưới đây là cách ML Toolkit có thể sử dụng hàm sigmoid để phân tích và phát hiện tấn công:

❖ Công thức hàm sigmoid:

$$\mathbf{f(z) = \frac{1}{1+e^{-z}}} \quad \mathbf{(1)}$$

Trong đó:

- **f(z)**: Đầu ra trong khoảng từ [0,1](giá trị xác suất ước lượng) (ví dụ: tấn công DoS).
- e: Hằng số Euler, cơ số của log tự nhiên ($e = \approx 2.71828$).
- **z**:Giá trị tuyến tính được tính toán từ dữ liệu đầu vào.

$$\mathbf{z=w^T x+b} \quad \mathbf{(2)}$$

Trong đó :

- **w**: Vector trọng số đại diện cho mức độ ảnh hưởng của từng đặc trưng x_i đến kết quả đầu ra.
- **x**: Vector dữ liệu đầu vào, gồm các đặc trưng $x_1, x_2, x_3, \ldots..$
- $\boldsymbol{w^T}$x: Là tích vô hướng giữa vector trọng số w và vector đặc trưng x.(w^Tx = $w_1x_1+w_2x_2+\cdots w_nx_n$).
- **b**: Hệ số chặn (bias), cho phép mô hình dịch chuyển giá trị đầu ra, để không bắt buộc đường hồi quy phải đi qua gốc tọa độ.

Hàm sigmoid có dạng một đường cong hình chữ S. Đóng vai trò là hàm liên kết ánh xạ sự kết hợp tuyến tính của các features đầu vào thành một xác suất. Hàm sigmoid đối xứng quanh điểm giữa của nó tại x=0.5(0.5 là giá trị ngưỡng).

❖ Quy trình ứng dụng sigmoid trong MLTK để phát hiện tấn công DoS

Bước 1: Thu thập và xử lý dữ liệu mạng

Dữ liệu đầu vào (x) có thể bao gồm các đặc trưng liên quan đến hoạt động mạng:

- Số lượng gói tin gửi mỗi giây (Packets per second).
- Lượng băng thông sử dụng (Bandwidth usage).
- Số lượng yêu cầu đồng thời (Concurrent requests).
- Tần suất gửi các yêu cầu bất thường.

Ví dụ: Một mẫu dữ liệu thu thập từ một sự kiện mạng có thể như sau: x = [150 packets/s, 500 KB/s bandwidth, 1000 requests/s]

Bước 2: Mô hình hóa xác suất bằng logistic regression

Logistic regression kết hợp các đặc trưng đầu vào với trọng số (w) được học qua quá trình huấn luyện, tính ra giá trị z:

$$\mathbf{z = w1.x1 + w2.x2 + w3.x3 + b}$$

Ví dụ: Với w = [0.2, 0.5, 0.3], b= -0.1, dữ liệu đầu vào là x = [150, 500, 1000]:

$$z= 0{,}2150+0{,}5500+0{,}31000 -0{,}1 = 440$$

Giá trị 2 được đưa qua hàm sigmoid để tính xác suất P(y = 1|x):

$$f(z) = \frac{1}{1+e^{-z}} \text{ Với z = 440: } f(z) = \frac{1}{1+e^{-440}} \approx 1$$

Kết quả: Xác suất rất cao. gần bằng 1. cho thấy đây là một sự kiên bất thường.

Bước 3: Phân loại (Classification)

Sau khi tính được xác suất f(z), Splunk chọn một ngưỡng (threshold) để quyết định:

- Nếu f(z) ≥ 0.5: Sự kiện được phân loại là bất thường (anomalous).
- Nếu f(z) < 0.5: Sự kiện được phân loại là bình thường (normal).

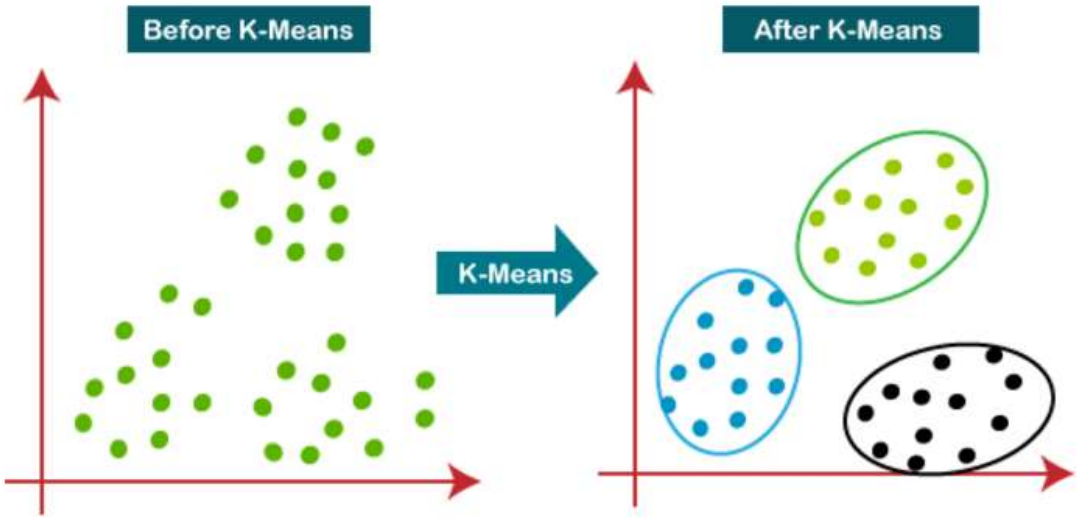
b. Thuật toán K-Means

K-means là một thuật toán phân cụm không giám sát phổ biến trong học máy, được sử dụng để phân nhóm các đối tượng trong một tập dữ liệu sao cho các đối tượng trong cùng một nhóm (cluster) có sự tương đồng cao và khác biệt với các nhóm khác. Mục tiêu của thuật toán K-means là chia tập dữ liệu thành K cụm, trong đó mỗi cụm có một trung tâm ,và các đối tượng trong cụm được phân loại dựa trên khoảng cách đến trung tâm của cụm đó.

Trong K-means để đánh giá mức độ giống nhau hay khoảng cách giữa 2 điểm dữ liệu ta có thể sử dụng các phép đo khoảng cách khác nhau :

$$(\boldsymbol{d(X,Y) = \sqrt{((x_1 + y_1)^2 + (x_2 + y_2)^2 + \ldots. + (x_n + y_n)^2}}) \textbf{ (3)}$$

Ngoài khoảng cách Euclidean, tuỳ thuộc vào từng bài toán có thể sử dụng phương pháp đo khác (cosine, manhattan…).



Hình 3.5: Biểu đồ biểu diễn K-Means

3.2.4 Machine Learning Toolkit Splunk

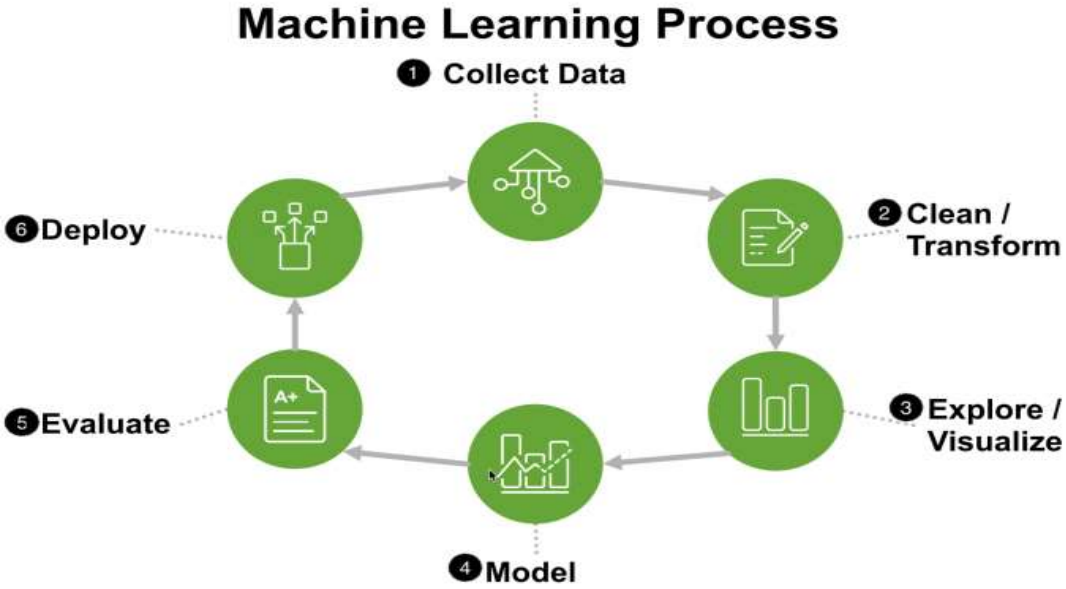
Machine Learning Toolkit (MLTK) trong Splunk là một công cụ mạnh mẽ cho phép người dùng áp dụng các thuật toán học máy vào dữ liệu được thu thập và lưu trữ trong Splunk. MLTK cung cấp giao diện đơn giản và trực quan để tạo, huấn luyện, đánh giá, và triển khai các mô hình học máy mà không yêu cầu kiến thức lập trình chuyên sâu. Người dùng có thể sử dụng MLTK để phát hiện bất thường, dự báo xu hướng, phân tích hành vi, và tự động hóa các quy trình dựa trên dữ liệu. Với khả năng tích hợp chặt chẽ trong Splunk, MLTK giúp doanh nghiệp tận dụng dữ liệu để đưa ra các quyết định thông minh và dự đoán chính xác.

Công nghệ máy học (Machine Learning - ML) đóng một vai trò quan trọng trong lĩnh vực an ninh mạng bằng cách cung cấp khả năng phát hiện và phản ứng tự động đối với các mối đe dọa mạng. Dưới đây là hai ứng dụng chính của máy học trong an ninh mạng:

- Phát Hiện Hành Vi Bất Thường: MLTK có thể phân tích dữ liệu từ các hoạt động mạng để xác định các hành vi không bình thường. Bằng cách học từ dữ liệu lịch sử, các thuật toán máy học có thể nhận diện các mẫu không bình

thường, bao gồm các cuộc tấn công hoặc hoạt động đáng ngờ trên mạng.Hỗ trợ tốt trong việc phát hiện các cuộc tấn công bất ngờ nhắm vào hệ thống .

- Phân Tích và Dự Đoán Tấn Công: Máy học có thể được sử dụng để phân tích dữ liệu mạng và dự đoán các cuộc tấn công sắp xảy ra. Bằng cách xây dựng các mô hình dự đoán từ dữ liệu lịch sử và các thông tin từ các mối đe dọa mạng hiện tại, máy học có thể cung cấp cái nhìn về các cuộc tấn công có thể xảy ra trong tương lai.Điều này giúp cho người dùng hoặc doanh nghiệp đưa ra các biện pháp phòng ngừa để hạn chế mức độ thiệt hại đến từ các cuộc tấn công .



Hình 3.6: Quy trình học máy

Quy trình học máy:

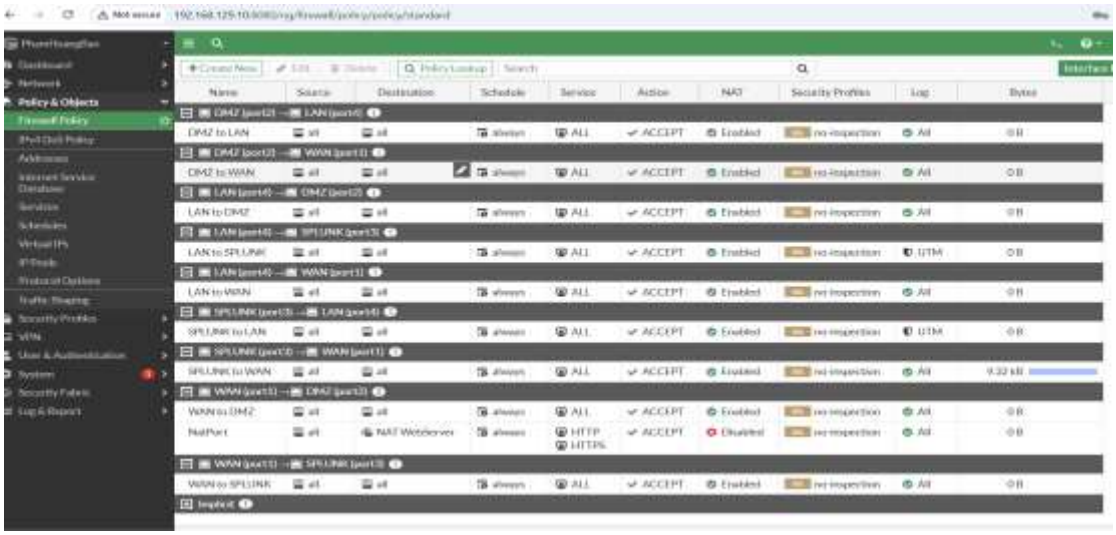
- ❖ Thu thập dữ liệu có sẵn như phần trăm CPU, mức sử dụng bộ nhớ, nhiệt độ máy chủ, dung lượng ổ đĩa hoặc giá trị bán hàng.
- ❖ Làm sạch và chuyển đổi dữ liệu đó. Tất cả máy học đều mong đợi một ma trận số làm đầu vào. Nếu bạn đang thu thập dữ liệu thiếu giá trị, thì bạn cần phải làm sạch và chuyển đổi dữ liệu đó cho đến khi nó ở dạng mà máy học yêu cầu.
- ❖ Khám phá và trực quan hóa dữ liệu để đảm bảo dữ liệu đang mã hóa đúng những gì bạn mong đợi.
- ❖ Xây dựng mô hình trên dữ liệu đào tạo.
- ❖ Đánh giá dữ liệu thử nghiệm mô hình.
- ❖ Triển khai mô hình trên dữ liệu chưa biết.

3.3 Quá trình thực nghiệm

3.3.1 Cài đặt và thiết lập hệ thống

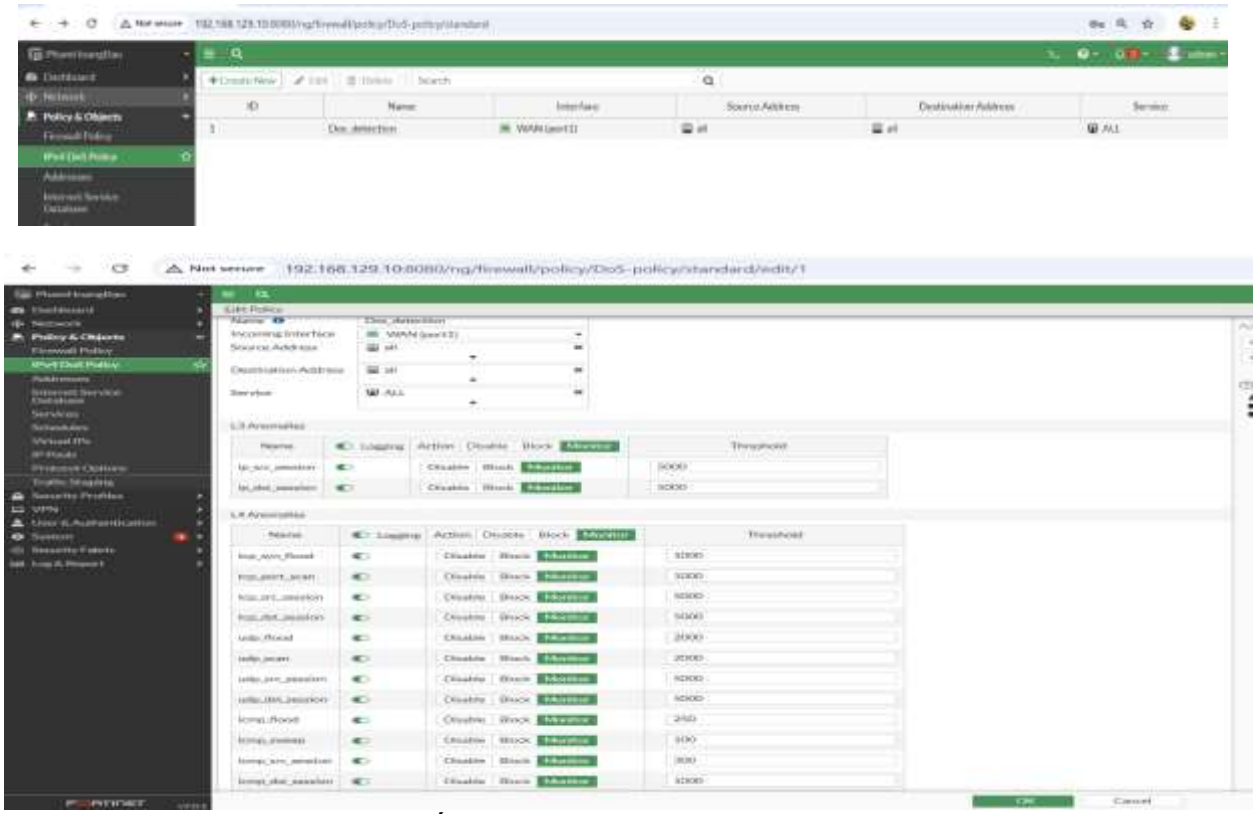
❖ **Firewall (Fortigate)**

- Cấu hình Policies cho Firewall



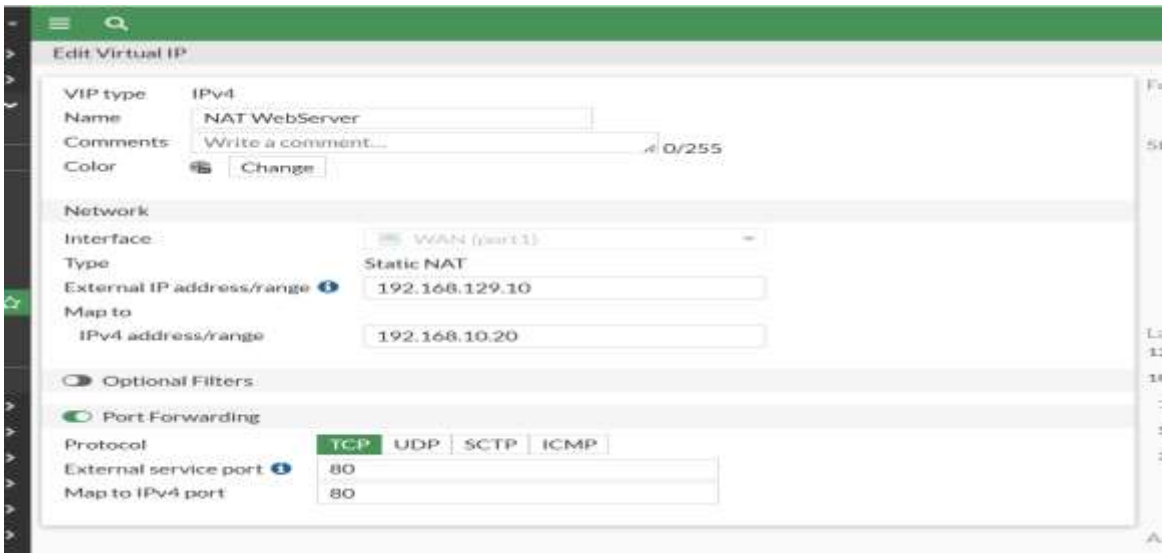
Hình 3.7 : Các policy cho từng vùng trong mô hình mạng

- Thiết lập các ngưỡng cảnh báo để chặn bằng IPv4DoS Policy của Fortigate



Hình 3.8: Thiết lập các ngưỡng cảnh báo cho Firewall Fortigate

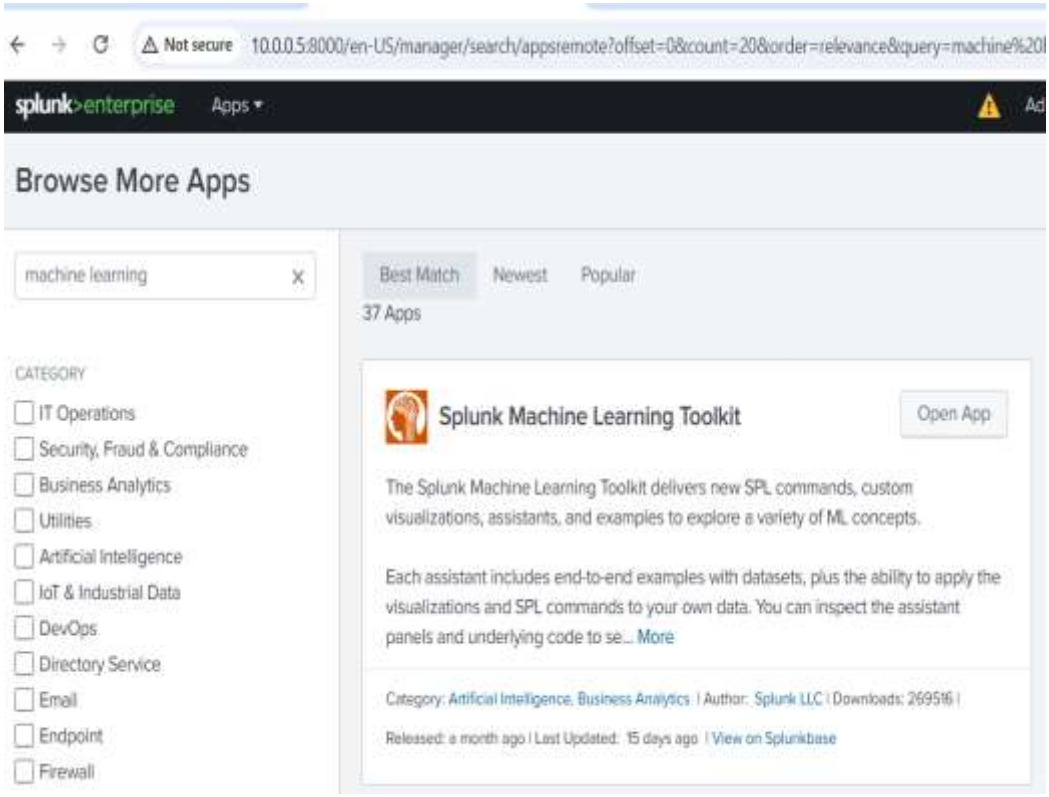
- Cấu hình Nat Port Web Server



Hình 3.9 :Cấu hình Virtual IP

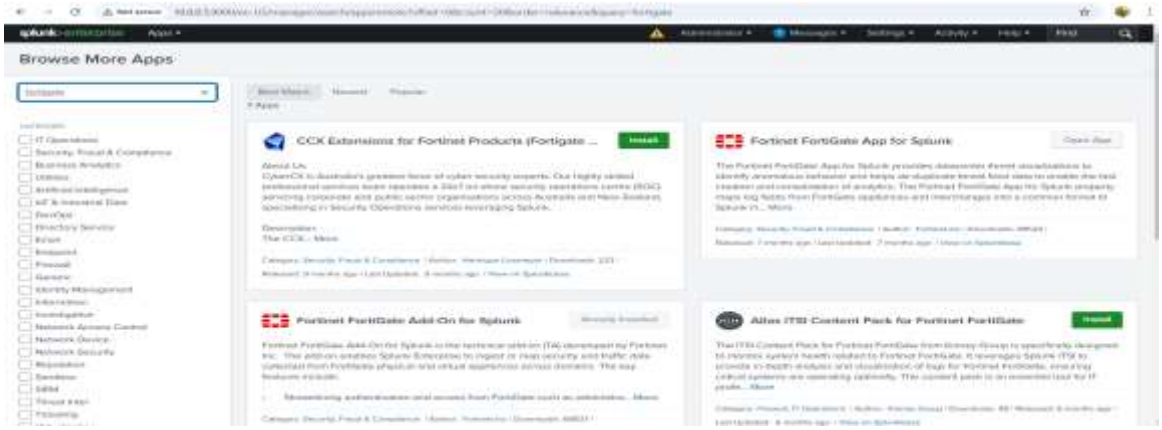
❖ **Splunk**

- Cài đặt Machine Learning Toolkit trong Splunk



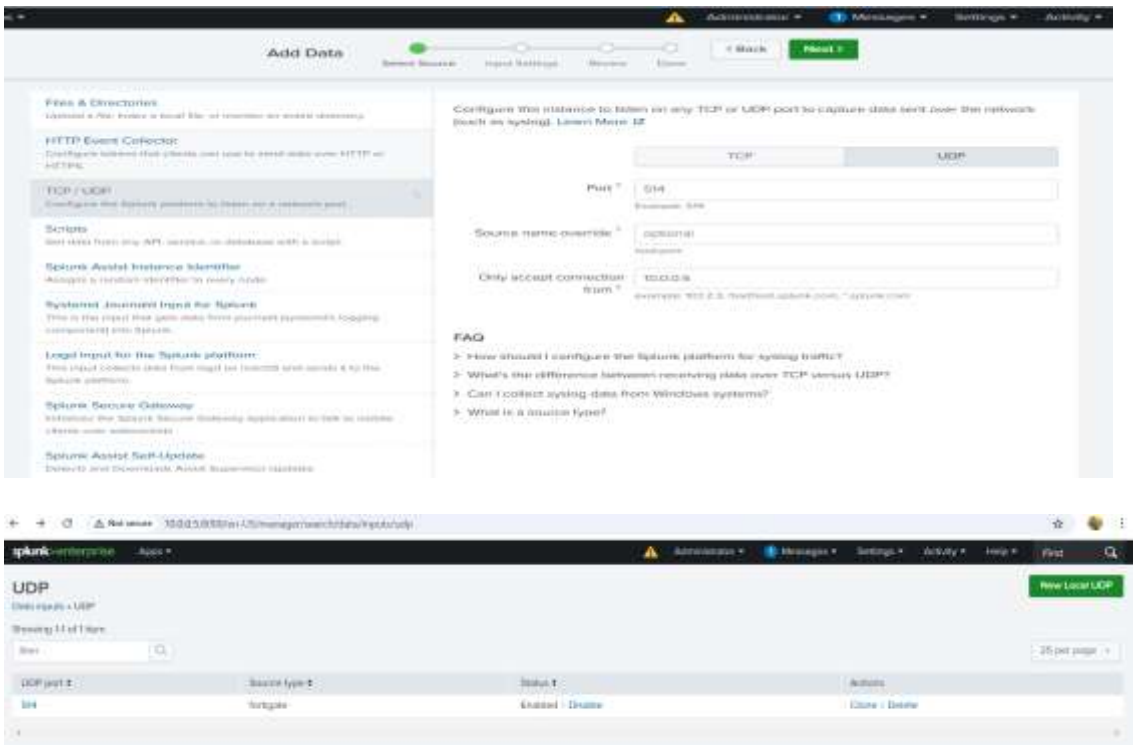
Hình 3.10: Splunk Machine Learning Toolkit

- Cài đặt Fortinet FortiGate App for Splunk và Fortinet FortiGate App-On for Splunk để nhận log từ Fortigate



Hình 3.11 : Ứng dụng hỗ trợ từ Fortinet trên Splunk.

- Cấu hình để nhận log từ Fortigate thông qua port 514 bằng giao thức UDP



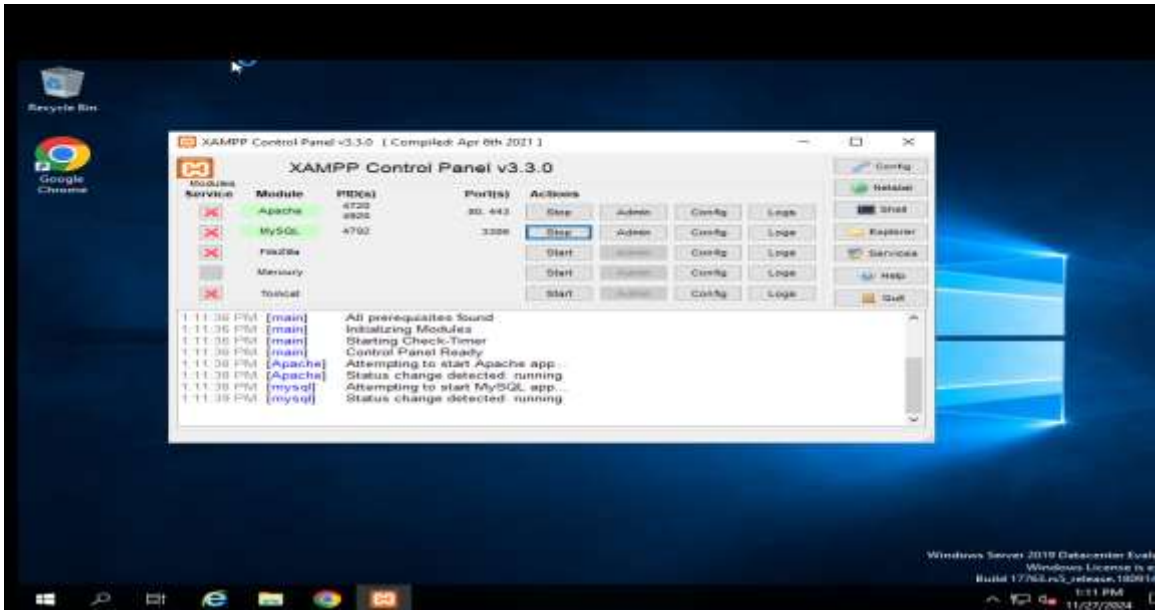
Hình 3.12:Tạo kết nối đến fortigate

- Tạo Index cho Fortigate

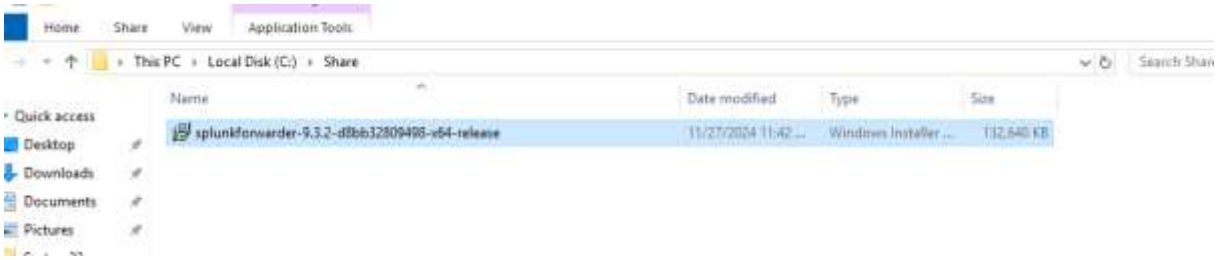
Hình 3.13: Tạo 1 Index để chứa các log từ splunk

❖ **Web Server (Windows Server 2019)**

- Sử dụng công tụ Xampp để khởi chạy website và Cài đặt Splunk Forwarder, kết nối đến Splunk Server.



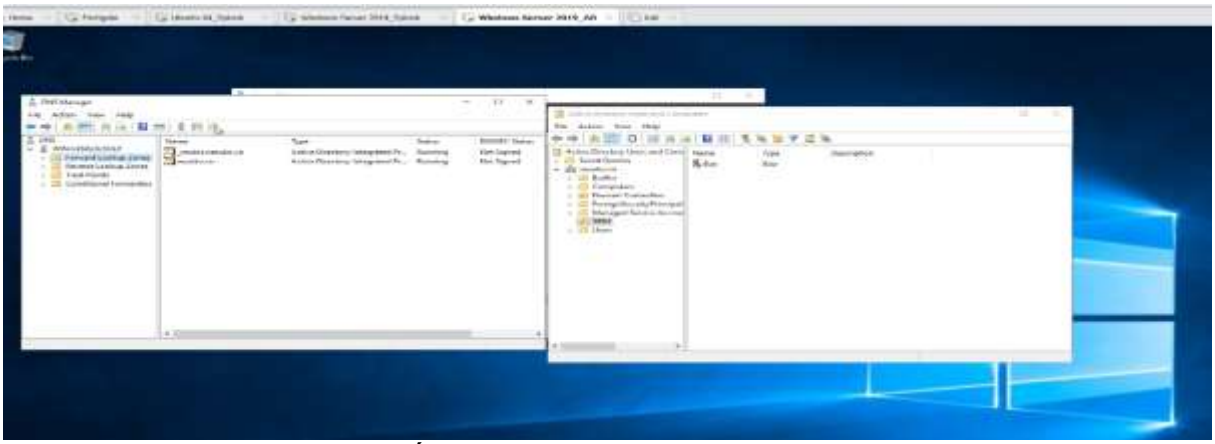
Hình 3.14: Sử dụng Xampp để mở port 80,443 để chạy website



Hình 3.15: Phần mềm Splunk Forwarder cài đặt tại máy webserver.

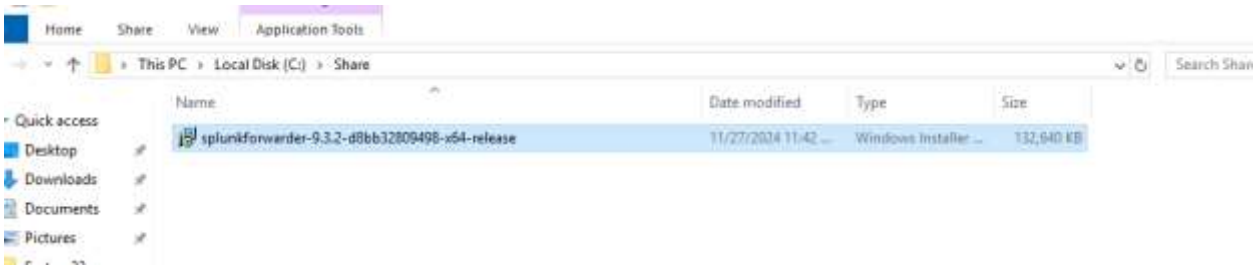
❖ **AD Server (Windows Server 2019)**

- Cài đặt Domain
- Tạo User và Domain Name



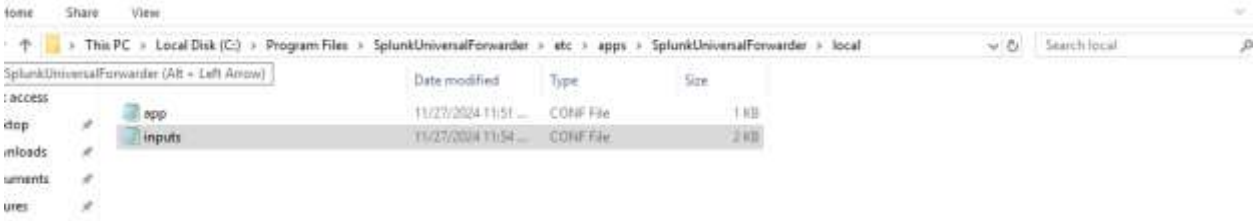
Hình 3.16: Thiết lập các user cho các nhân viên tại vùng mạng Lan

- Cài đặt Splunk Forwarder, kết nối đến Splunk Server
- Tải file setup trên trên chủ về máy



Hình 3.17: Splunk Forwarder cài đặt tại máy server AD.

- Vào ổ đĩa C (nơi lưu cài đặt Splunk Forward)
- ProgramFile->SplunkUniversalForwarder->etc->apps->SplunkUniversalForwarder -> local để chỉnh sửa file input

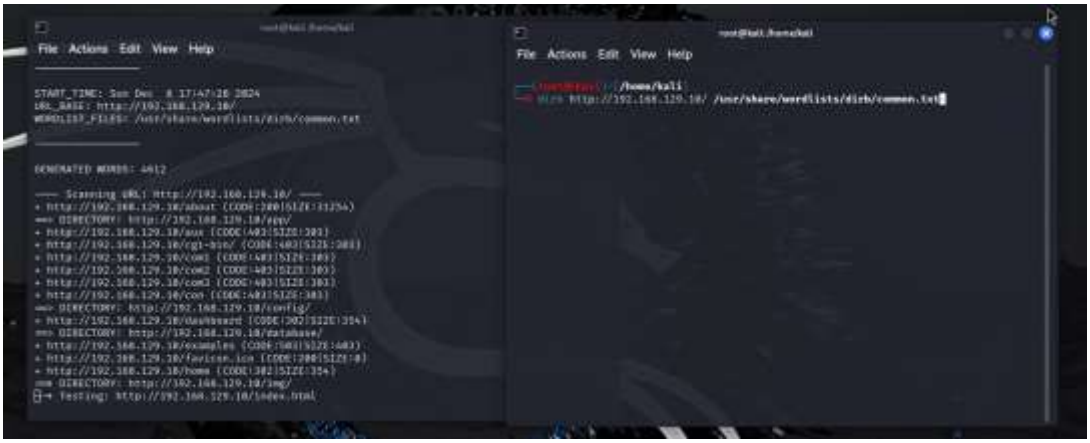


Hình 3.18: File input để gửi log đến splunk

- Thêm index=ad_log ,Ad_log là index đã tạo trong splunk để nhận log từ máy Domain .Vào file bin của Splunk Foward, khởi động lại Splunk Forward bằng CMD

3.3.2 Kịch bản thử nghiệm phát hiện tấn công Brute Froce (Directory Brute Forcing)

- Thực hiện tấn công Directory Brute Forcing (công cụ dirbuster) quét các đường dẫn của website , dùng Splunk để phát hiện và đưa ra giải pháp ngăn chặn.
- Máy Attacker có địa chỉ là 192.168.129.134, sử dụng công cụ dirbuster để tấn công Brute Force vào WebServer.



Hình 3.19 :Sử dụng công cụ dirbuster.

- Phát hiện log Apache, các yêu cầu POST tới /login với IP 192.168.129.133 thực hiện 16 yêu cầu trong 1 phút.

Hình 3.20:Truy cập bất thường đến các đường dẫn không tồn tại của website.

Giải thích truy vấn :

- **index=webserver_log sourcetype=apache:error**
  - Lọc dữ liệu từ index webserver_log, và chỉ lấy dữ liệu từ nguồn apache:error.
- **| where _time >= relative_time(now(), "-10m")**
  - Áp dụng điều kiện để chỉ lấy các truy vấn đến website trong vòng 10 phút gần nhất.
- **| search "Forbidden" OR "doesn't point to a file or directory"**
  - Lọc lấy các bản ghi chứa chuỗi "Forbidden" hoặc "doesn't point to a file or directory" để tìm các lỗi liên quan đến quyền truy cập hoặc các yêu cầu tới các file hoặc thư mục không tồn tại.
- **| stats count by srcip_client**

- Thống kê số lần sự kiện xuất hiện, nhóm theo IP của client để xác định tần suất các lỗi xảy ra từ từng IP cụ thể.

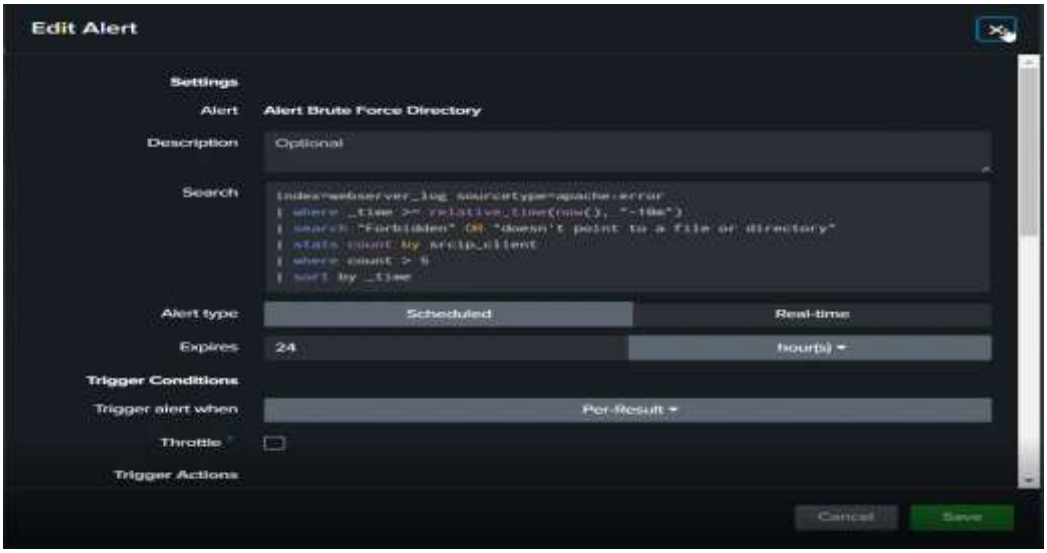
❖ **| where count > 5**

- Điều kiện lọc để chỉ giữ các IP client có số lần xuất hiện của lỗi lớn hơn 5.

❖ **| sort by _time**

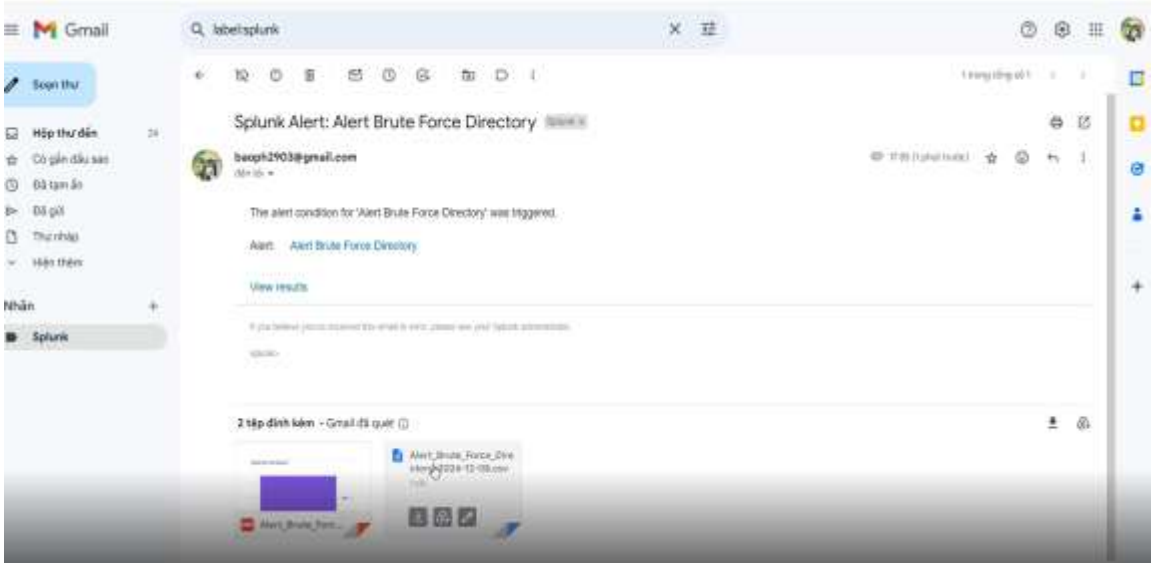
- Sắp xếp kết quả theo thứ tự thời gian.

- Cấu hình cảnh báo khi phát hiện tấn công Directory Brute Forcing.



Hình 3.21: Alert khi có tấn công Brute Force

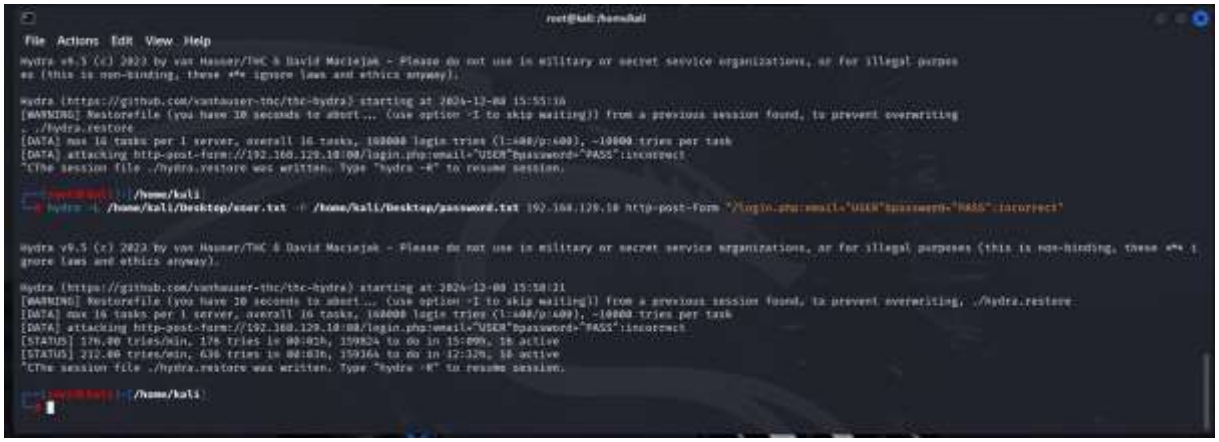
- Thông báo được gửi về khi phát hiện tấn công Directory Brute Forcing.



Hình 3.22: Mail alert Directory Brute Forcing được gửi về

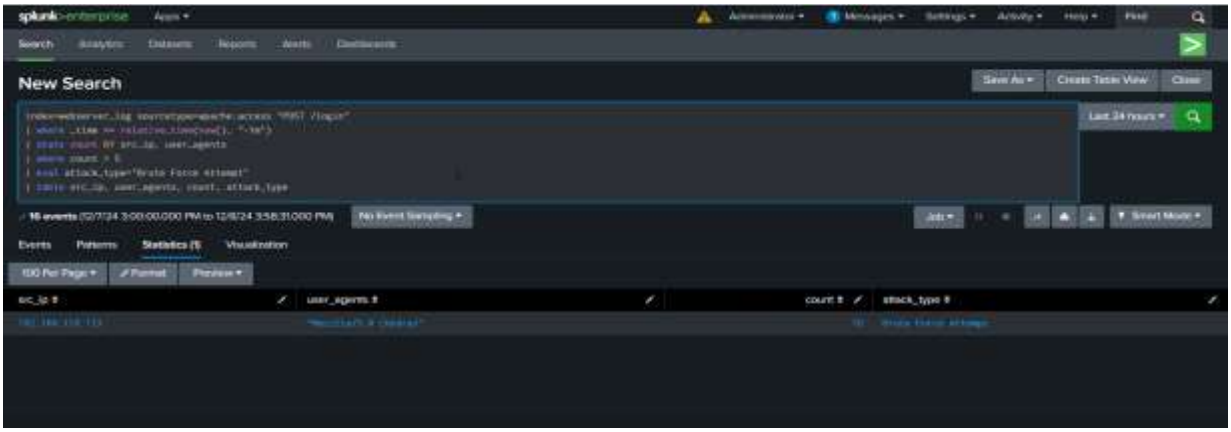
3.3.3 Kịch bản thử nghiệm phát hiện tấn công Brute Froce(Password Brute Forcing)

- Thực hiện tấn công Password Brute Forcing (công cụ Hydra) quét các đường dẫn của website , dùng Splunk để phát hiện và đưa ra giải pháp ngăn chặn.
- Máy Attacker có địa chỉ là 192.168.129.133, sử dụng công cụ hydra để tấn công Brute Force vào WebServer.



Hình 3.23: Sử dụng công cụ Hydra.

- Phát hiện IP 192.168.129.133 đã thực hiện 16 yêu cầu POST tới trang login của website trong vòng 1 phút.



Hình 3.24: Truy vấn phát hiện các lần đăng nhập sai liên tiếp

Giải thích truy vấn :

- ❖ **index=webserver_log sourcetype=apache:access "POST /login"**
  - Lọc dữ liệu từ index webserver_log, và chỉ lấy dữ liệu từ nguồn apache:access.Truy vấn chỉ tập trung vào các yêu cầu HTTP POST được gửi đến endpoint /login.

❖ **| where _time >= relative_time(now(), "-1m")**

- Áp dụng điều kiện để chỉ lấy các truy vấn đến website trong vòng 1 phút gần nhất.

❖ **| stats count BY src_ip, user_agents**

- Thống kê số lượng yêu cầu , nhóm theo địa chỉ IP nguồn và user agent để đếm số lần yêu cầu từ từng IP .

❖ **| where count > 5**

- Lọc chỉ giữ lại các IP có hơn 5 yêu cầu tới trong vòng 1 phút

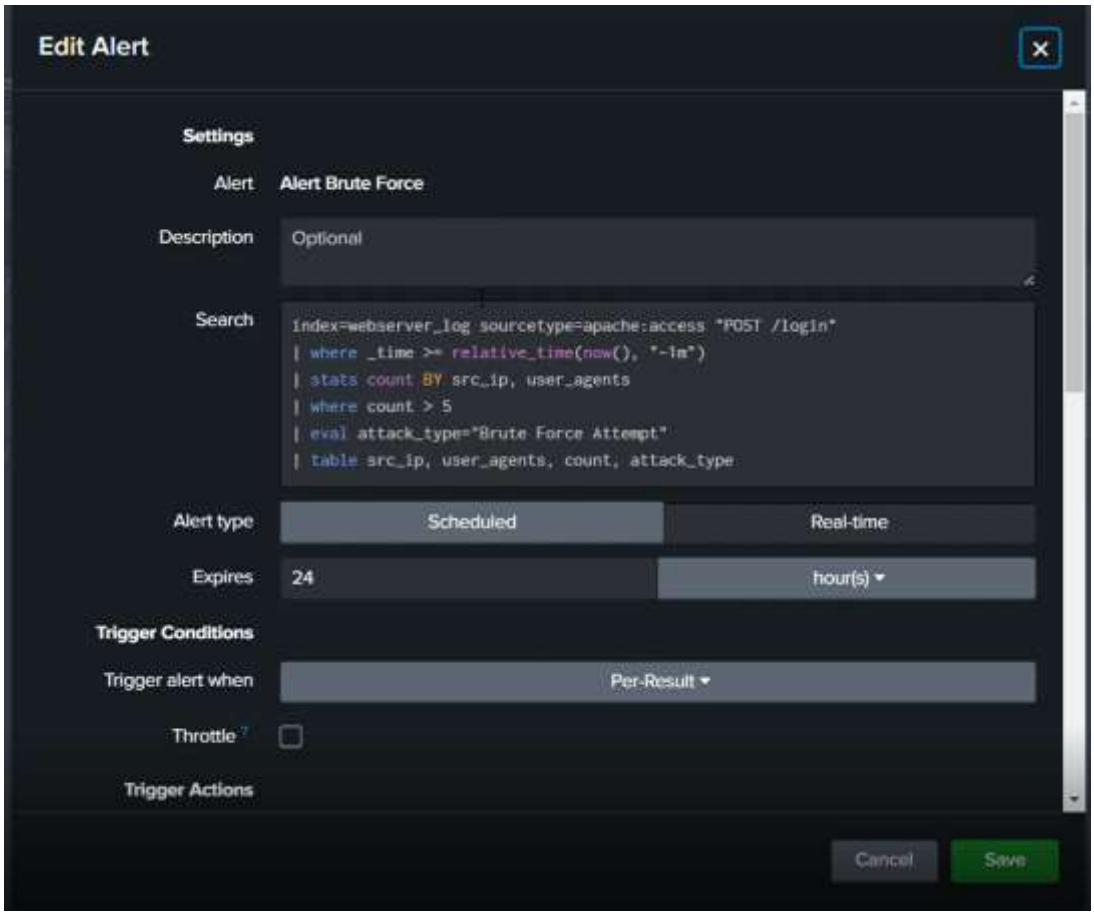
❖ **| eval attack_type="Brute Force Attempt"**

- Tạo một trường attack_type và gán giá trị "Brute Force Attempt" cho tất cả các kết quả.

❖ **| table src_ip, user_agents, count, attack_type**

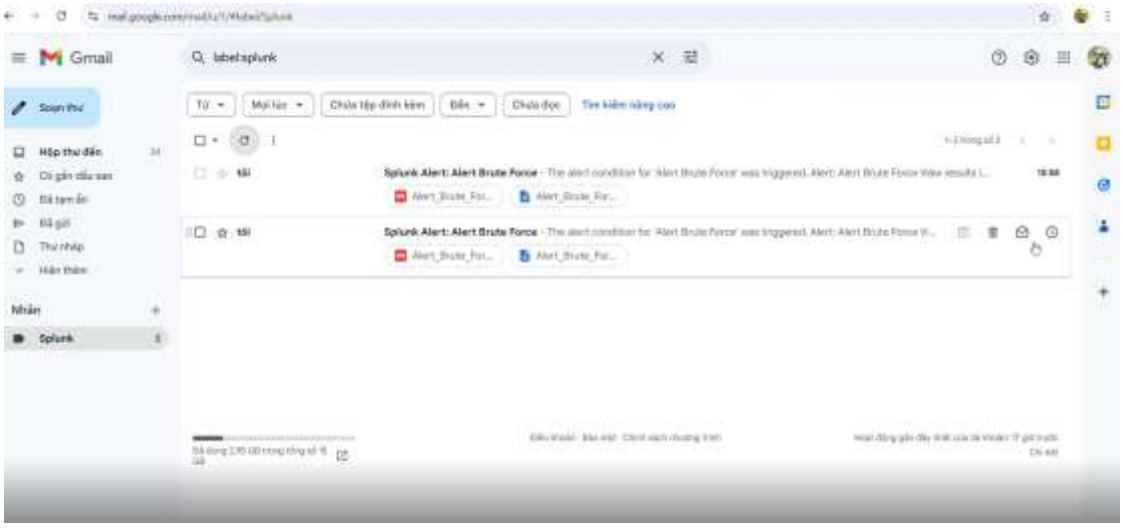
- Hiển thị kết quả theo các cột được chọn .

- Cấu hình cảnh báo khi phát hiện tấn công Password Brute Forcing.



Hình 3.25:Alert phát hiện khi có log đang nhập sai liên tục

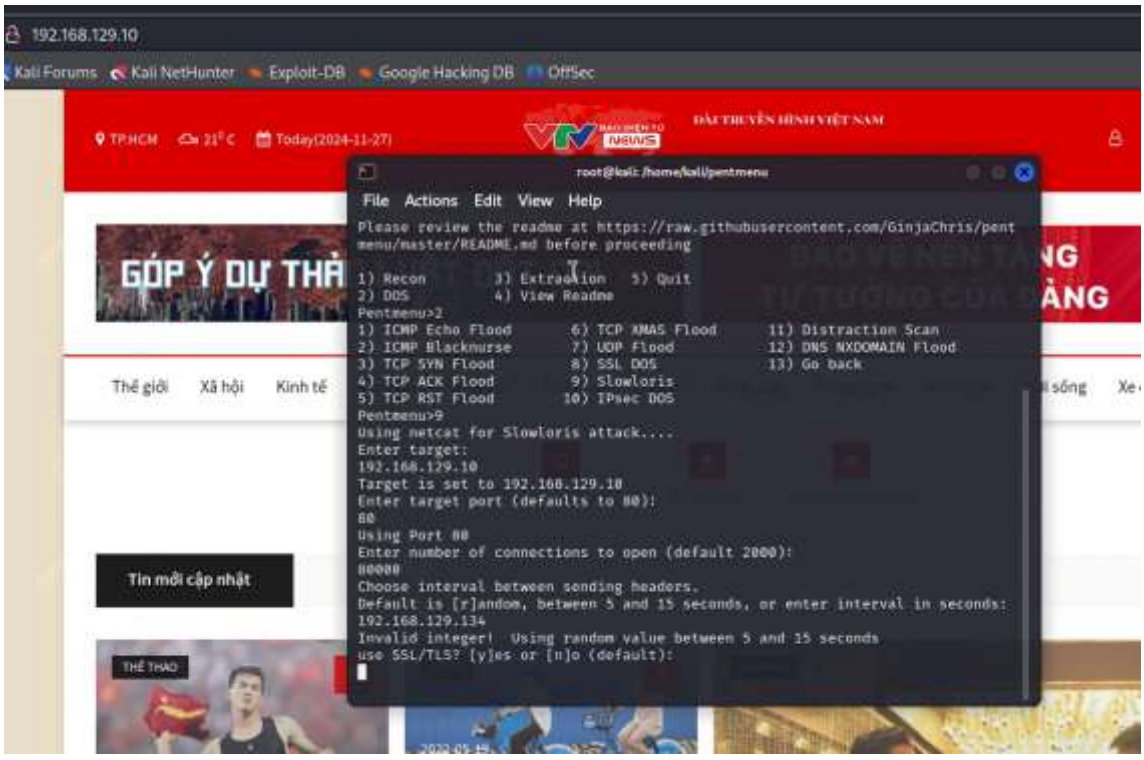
- Thông báo được gửi về khi phát hiện tấn công Password Brute Forcing.



Hình 3.26: Mail alert Password Brute Forcing được gửi về

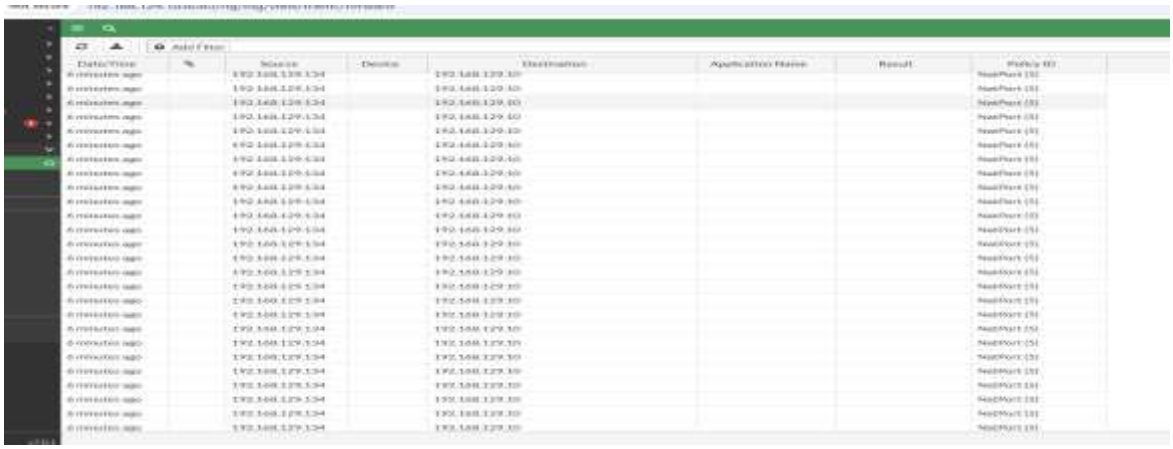
3.3.4 Kịch bản thử nghiệm phát hiện tấn công Dos

- Thực hiện tấn công DoS/DdoS(công cụ pentmenu) vào Webserver, dùng Firewall và Splunk để phát hiện và đưa giải pháp ngăn chặn.
- Máy Attacker có địa chỉ là 192.168.129.134, sử dụng công cụ Pentmenu để tấn công Dos vào WebServer



Hình 3.27 : Sử dụng pentmenu để tấn công Dos

- Phía tường lửa Fortigate nhận được rất nhiều Log từ máy 192.168.129.134 gửi yêu cầu đến máy 192.168.129.10

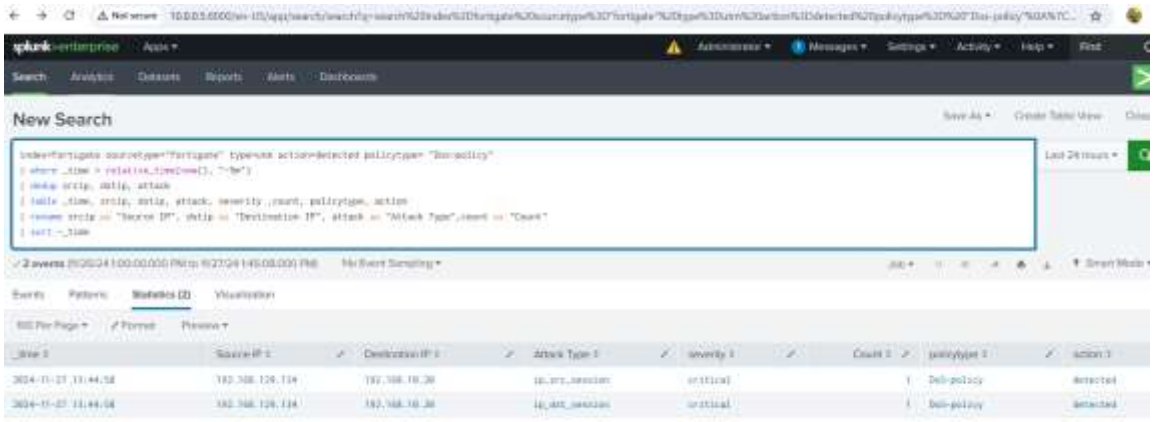


Hình 3.28: Log của firewall khi gặp Dos

- Mở tab Anomaly xem trang thái

Hình 3.29: Firewall thực hiện action detected khi gặp dos

- Về phía máy Splunk, sử dụng lệnh truy vấn, thì phát hiện cuộc tấn công Dos

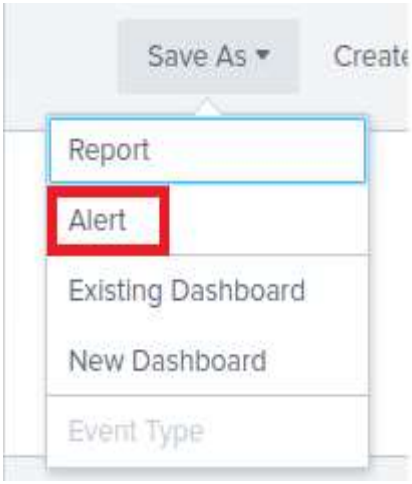


Hình 3.30: Truy vấn để phát hiện các log của Dos

Giải thích truy vấn:

- ❖ **index=fortigate sourcetype="fortigate" type=utm action=detected policytype="Dos-policy"**
  - Lọc dữ liệu trong chỉ mục "fortigate" với kiểu nguồn (sourcetype) là "fortigate", loại sự kiện là UTM, hành động là "detected", và chính sách loại "Dos-policy".
- ❖ **| where _time > relative_time(now(), "-5m") :**
  - Chỉ giữ các sự kiện xảy ra trong vòng 5 phút gần đây.
- ❖ **| dedup srcip, dstip, attack :**
  - Loại bỏ các bản ghi trùng lặp dựa trên IP nguồn, IP đích và loại tấn công.
- ❖ **| table _time, srcip, dstip, attack, severity ,count, policytype, action :**
  - Hiển thị các cột gồm thời gian, IP nguồn, IP đích, loại tấn công, mức độ nghiêm trọng, số lượng, loại chính sách và hành động.
- ❖ **| rename srcip as "Source IP", dstip as "Destination IP", attack as "Attack Type",count as "Count"**
  - Đổi tên các cột để rõ nghĩa hơn: srcip thành "Source IP", dstip thành "Destination IP", attack thành "Attack Type", và count thành "Count".
- ❖ **| sort -_time**
  - Sắp xếp các bản ghi theo thời gian (_time).

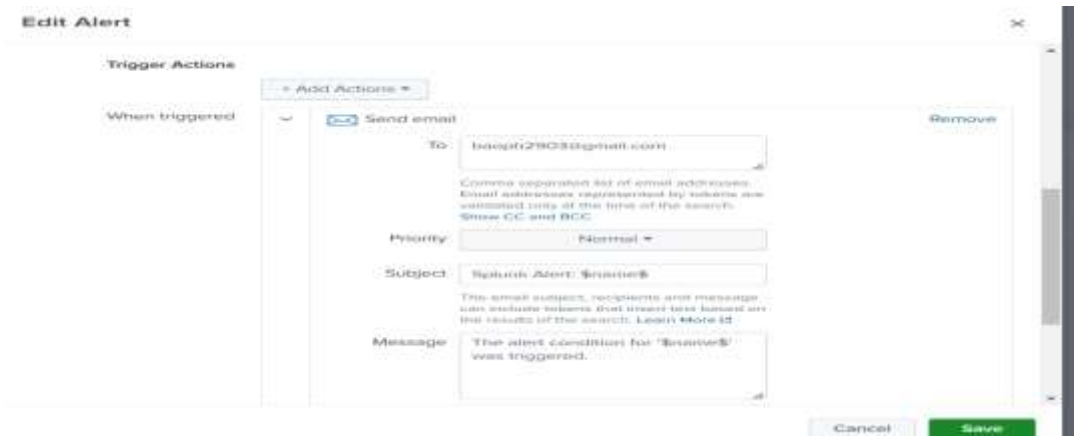
- Tiến hành cấu hình gửi Mail với câu truy vấn trên
- Save As -> Alert



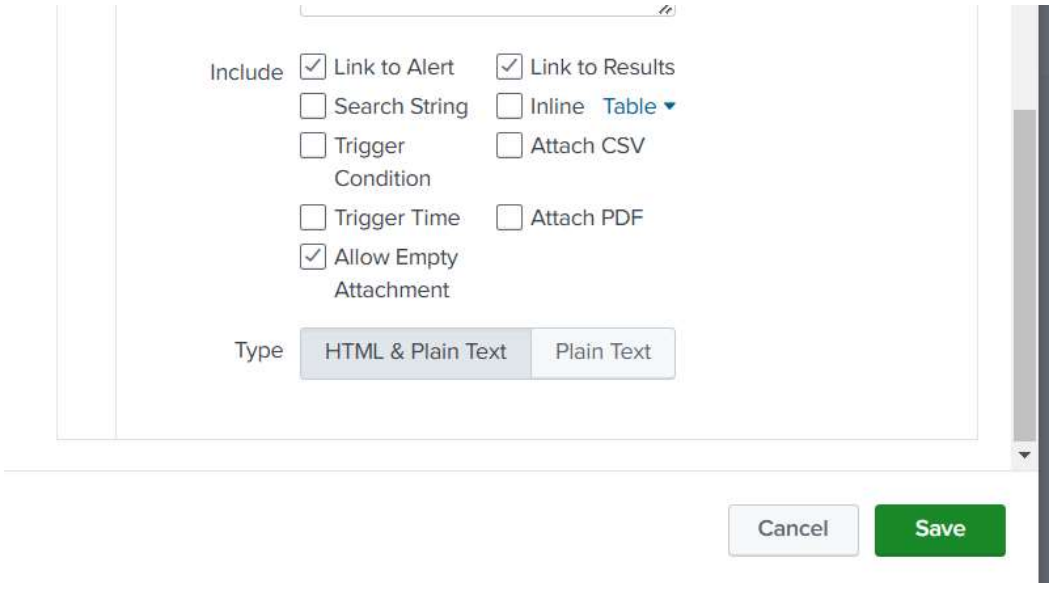
Hình 3.31: Tạo cảnh báo khi có log Dos

- Cấu hình cảnh báo khi nhận biết được lưu lượng Dos

Hình 3.32 : Lựa chọn cảnh báo theo lịch hoặc thời gian thực

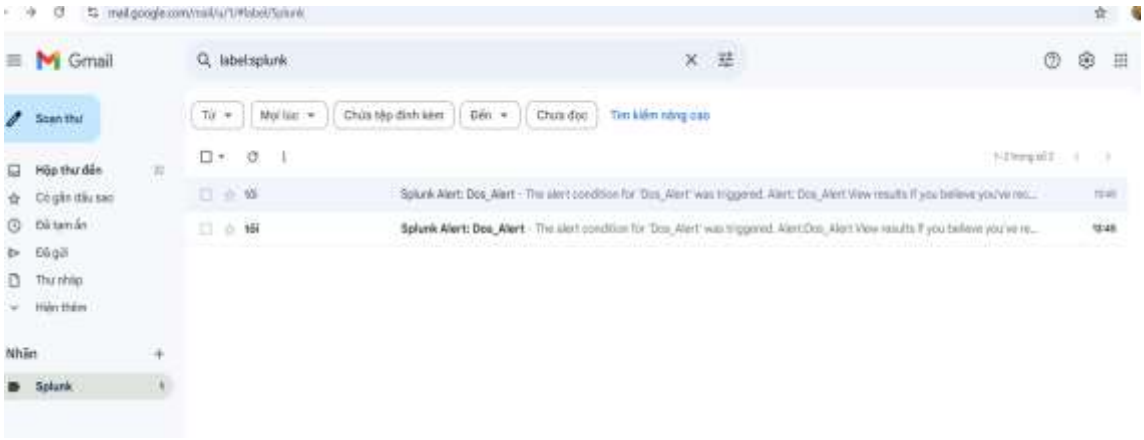


Hình 3.33: Thêm email của người quản trịnh để gửi cảnh báo



Hình 3.34: Tuỳ chọn loại tệp gửi cùng cảnh báo và định dạng cảnh báo

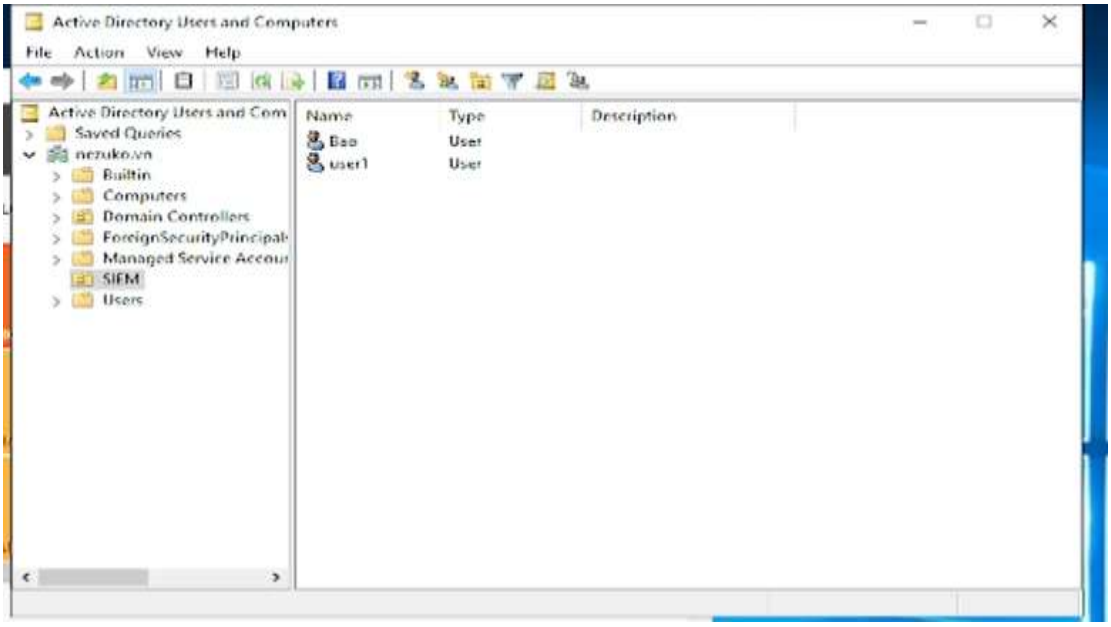
- Sau các bước thiết lập truy vấn và tạo cảnh bảo khi cuộc tấn công Dos diễn ra , hệ thống sẽ phát hiện và gửi cảnh báo đến Mail người quản trị.



Hình 3.35: Email cảnh báo được gửi từ splunk

3.3.5 Kịch bản thử nghiệm phát hiện thay đổi tại Server AD

- Thực hiện thay đổi các tính năng (thêm, xóa user trong AD DC). Dùng Splunk để phát hiện, phân tích và cảnh báo.
- Phía máy Domain Controller tạo user 1



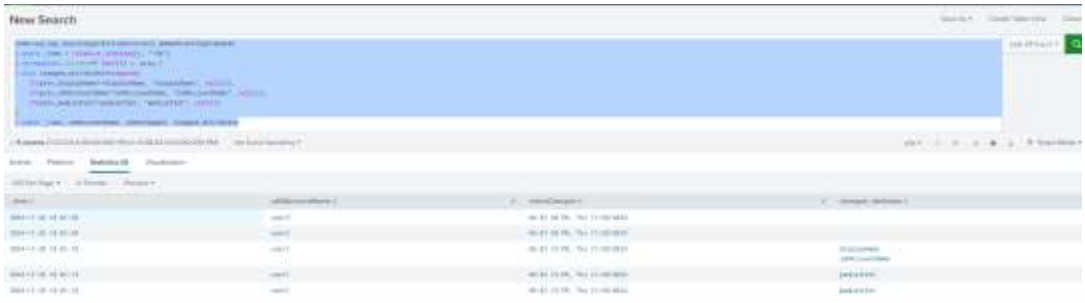
Hình 3.36: Bảng quản lý tài khoản user

- Trên máy Splunk, sử dụng truy vấn để thu thập thông tin log sự kiện từ máy DC

Giải thích truy vấn:

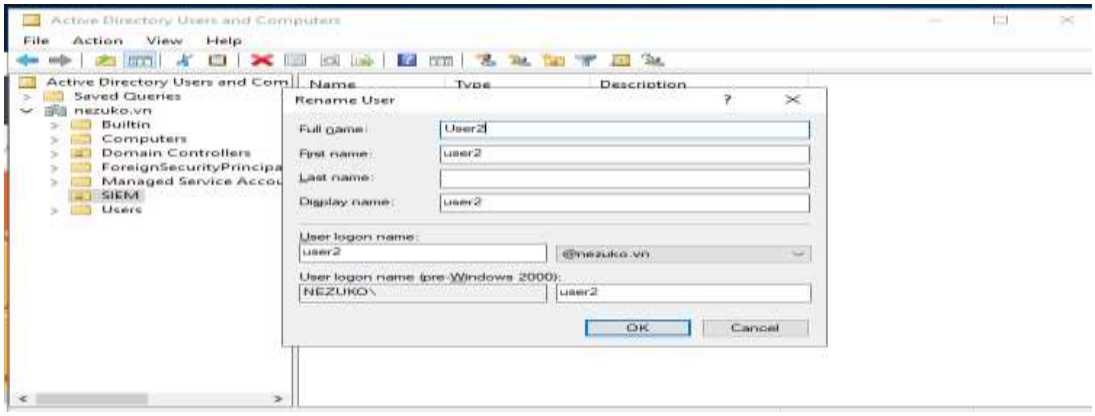
- **index=ad_log sourcetype=ActiveDirectory admonEventType=Update**
  - Truy vấn tìm kiếm trong index ad_log với loại dữ liệu từ Active Directory, lọc lấy những sự kiện có hành động "Update" (cập nhật có liên quan đến các thay đổi trong Active Directory).
- **| where _time > relative_time(now(), "-5m")**
  - Lọc các sự kiện chỉ xảy ra trong 5 phút gần nhất.
- **| streamstats current=f last(*) as prev_***
  - Sử dụng streamstats để truy xuất các giá trị của tất cả các trường trong dòng trước đó và tạo các trường mới có tiền tố prev_. Điều này cho phép so sánh giá trị của các trường hiện tại với giá trị của chúng trong các sự kiện trước đó.
- **| eval changed_attributes=mvappend(...)**
  - Tạo một trường mới, changed_attributes, chứa danh sách các thuộc tính đã thay đổi.
- **| if(prev_displayName!=displayName, "displayName", null()):**
  - So sánh giá trị hiện tại của displayName với giá trị trước đó (prev_displayName). Nếu khác nhau, thêm "displayName" vào danh sách thay đổi, nếu không thì bỏ qua .
  - Tương tự, các thuộc tính khác như sAMAccountName và pwdLastSet cũng được so sánh và thêm vào danh sách thay đổi nếu có sự khác biệt.
- **| table _time, sAMAccountName, whenChanged, changed_attributes**
  - Hiển thị kết quả dưới dạng bảng với các cột gồm thời gian của sự kiện (_time), tên tài khoản người dùng (sAMAccountName), thời gian thay đổi của tài khoản (whenChanged), và danh sách các thuộc tính đã thay đổi (changed_attributes).

- User1 đã được tạo lúc 18 giờ 1 phút



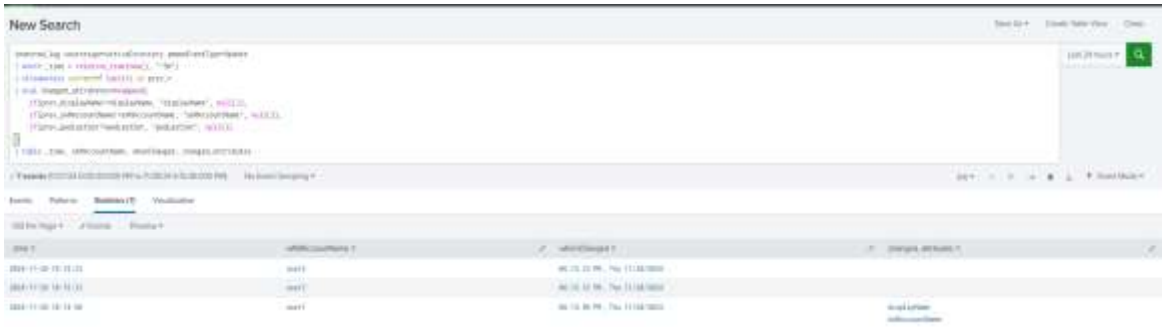
Hình 3.37:truy vấn thu thập thông tin log từ máy DC

- Trên máy DC, thay đổi username của User1 thành User2



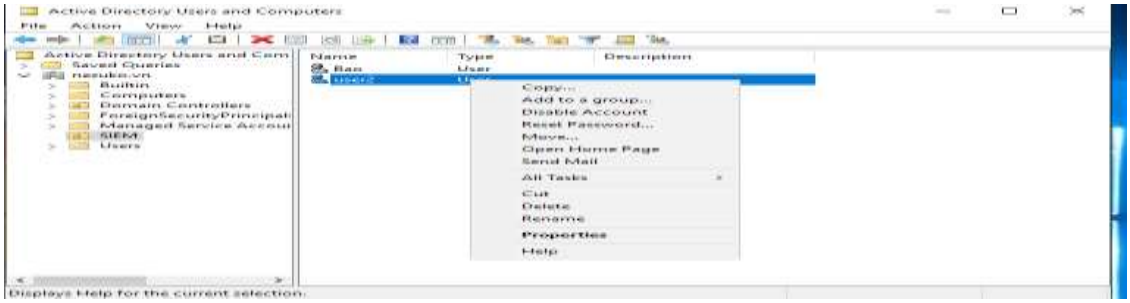
Hình 3.38:Bảng thông tin user tại máy DC

- Cùng lúc đó bên phía máy Splunk cũng nhận được log sự kiện về sự thay đổi này.



Hinh 3.39: Log của máy DC gửi đến splunk server.

- Máy DC xóa người dùng User2



Hình 3.40: Bảng quản lý user tại máy DC.

- Bên phía máy Splunk, sử dụng truy vấn để theo dõi sự kiện log.

Giải truy vấn:

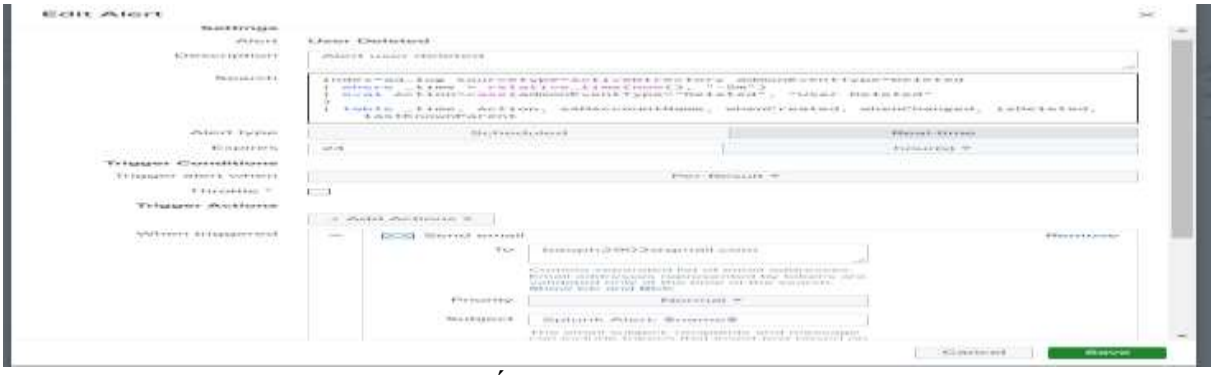
- **index=ad_log sourcetype=ActiveDirectory admonEventType=Deleted**
  - Tìm kiếm trong index ad_log với nguồn ActiveDirectory và lọc các sự kiện có kiểu hành động Deleted (các sự kiện xóa tài khoản người dùng trong Active Directory).
- **| where _time > relative_time(now(), "-5m")**
  - Lọc các sự kiện xảy ra trong 5 phút gần đây.
- **| eval Action=case(admonEventType="Deleted", "User Deleted")**
  - Tạo một trường mới có tên Action, gán giá trị "User Deleted" nếu kiểu sự kiện là "Deleted", để giúp xác định rõ ràng hành động xóa người dùng.
- **|table _time, Action, sAMAccountName, whenCreated, whenChanged, isDeleted, lastKnownParent**
  - Hiển thị kết quả dưới dạng bảng với các giá trị theo cột: thời gian sự kiện xóa (_time), hành động (Action), tên đăng nhập người dùng (sAMAccountName), thời gian tạo và thay đổi tài khoản (whenCreated, whenChanged), trạng thái xóa (isDeleted), và đối tượng cha cuối cùng của người dùng (lastKnownParent).

Hình 3.41: Log của sự kiện xoá user2 .

- Vào Save -> Alert để tạo cảnh báo(tất cả cảnh báo ở đây sẽ được gửi về mail)

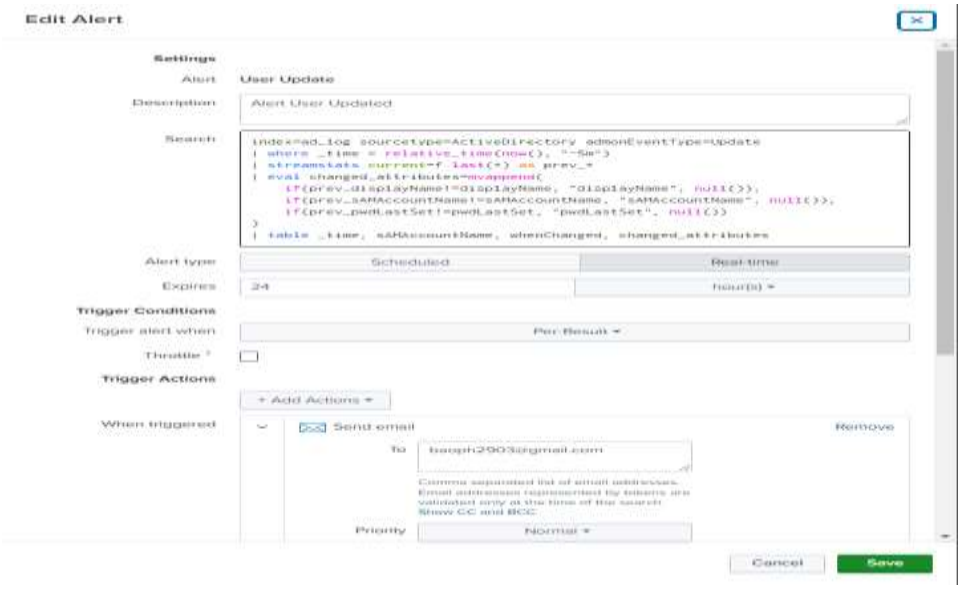
Hình 3.42: Tạo cảnh báo từ splunk.

- Cấu hình Alert cảnh báo Delete



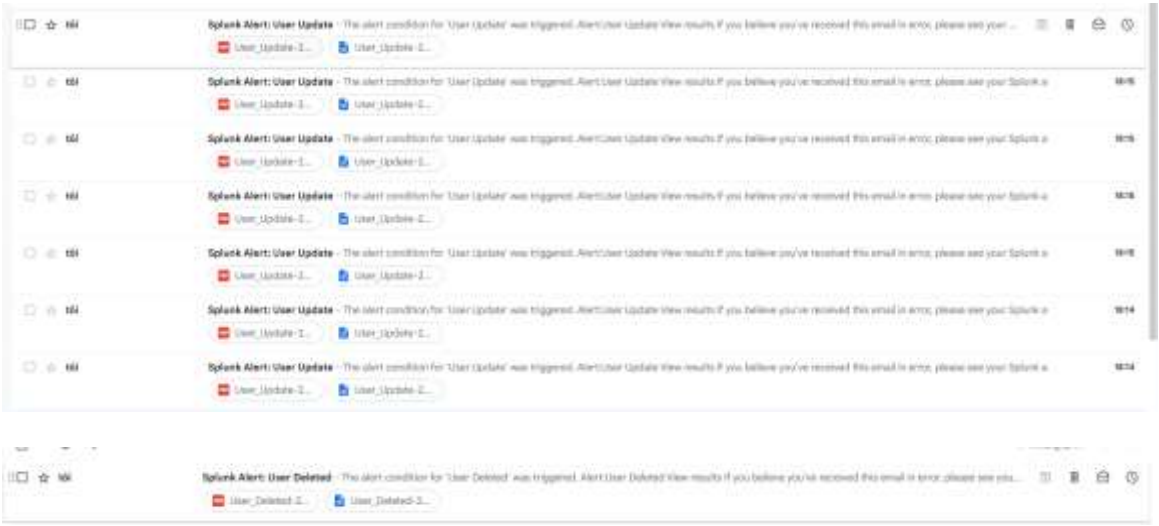
Hình 3.43: Thiết lập cảnh báo khi có sự kiện xoá user.

- Đây là cảnh báo về việc tạo user1 và rename user1 thành user2



Hình 3.44:Thiết lập cảnh báo khi có sự kiện thay đổi user.

- Đây là cảnh báo về việc tạo user1 và rename user1 thành user2



Hình 3.45: Các cảnh báo gửi về khi phát hiện hai sự kiện .

3.3.6 Kịch bản thử nghiệm phát hiện tấn công bằng Machine Learning.

Hình 3.46 :Truy vấn chuẩn hoá dữ liệu

Bước quan trọng trong quá trình sử dụng học máy chính là chuẩn hoá được các dữ liệu mà hệ thống chúng ta đã thu thập được ,các trường của log đóng vai trò là nguyên liệu để thực hiện quá trình học máy .Việc này sẽ tạo điều kiện cho việc sử dụng các thuật toán tối ưu và giảm thiểu được các sai soát từ dữ liệu không quan trọng .

Giải thích truy vấn:

- **index=fortigate sourcetype=fortigate**
  - Lọc log từ index fortigate và sourcetype=fortigate. Điều này chỉ lấy dữ liệu liên quan đến FortiGate.
- **|stats count AS total_sessions, sum(sentpkt) AS total_sent_packets, sum(rcvdpkt)AStotal_received_packets,sum(sentbyte) AS total_sent_bytes, sum(rcvdbyte) AS total_received_bytes BY _time, srcip, dstip, proto, service**

- **stats**:
  - **count**: Đếm tổng số phiên, lưu vào biến total_sessions.
  - **sum(sentpkt)**: Tổng số gói tin đã gửi, lưu vào total_sent_packets.
  - **sum(rcvdpkt)**: Tổng số gói tin đã nhận, lưu vào total_received_packets.
  - **sum(sentbyte)**: Tổng số byte đã gửi, lưu vào total_sent_bytes.
  - **sum(rcvdbyte)**: Tổng số byte đã nhận, lưu vào total_received_bytes.
- **BY**: Nhóm các giá trị thống kê theo _time (thời gian), srcip (địa chỉ IP nguồn), dstip (địa chỉ IP đích), proto (giao thức), và service.

❖ **| eval total_traffic=total_sent_bytes + total_received_bytes**

- Tính tổng lưu lượng (total_traffic) bằng cách cộng tổng byte gửi và nhận.

❖ **|evalpacket_rate=(total_sent_packets+total_received_packets)/total_sessions**

- Tính tốc độ gói tin (packet_rate) bằng cách chia tổng số gói gửi và nhận cho tổng số phiên (total_sessions).

❖ **| table _time, srcip, dstip, total_sessions, total_traffic**

- Hiển thị các cột cụ thể: _time, srcip, dstip, total_sessions, total_traffic.

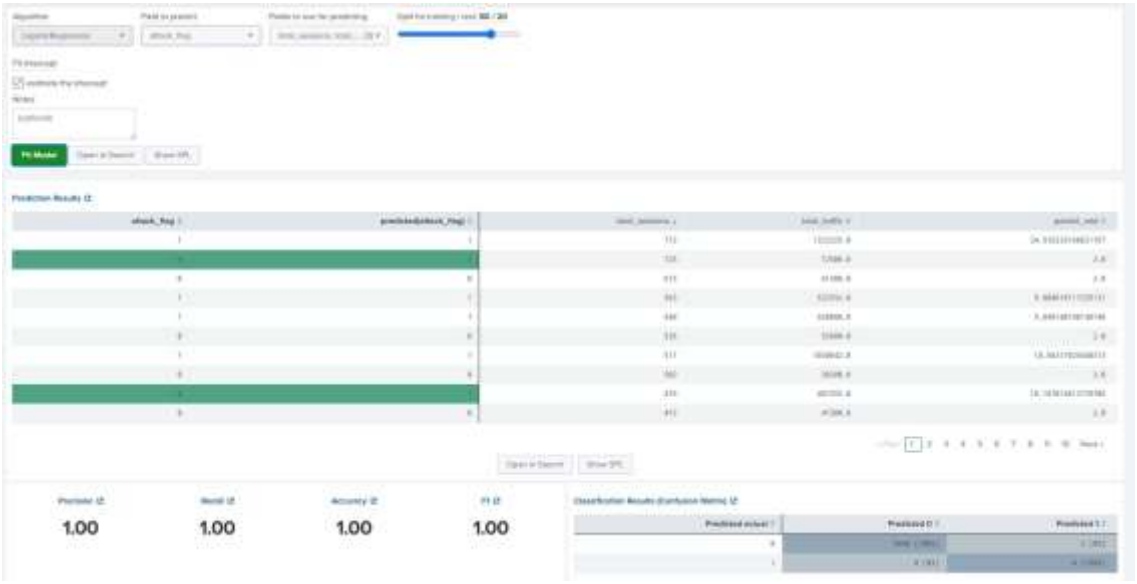
❖ **| eval attack_flag=if(total_sessions > 500 AND total_traffic > 100000 AND packet_rate > 10, 1, 0)**

- Thêm cột attack_flag để đánh dấu cảnh báo tấn công.

- **if**:
  - Nếu số phiên (**total_sessions**) lớn hơn 500 và tổng lưu lượng (**total_traffic**) lớn hơn 100,000 cộng với tốc độ gói tin(**packet_rate**) lớn hơn 9, thì đặt attack_flag là 1 (có khả năng tấn công).
  - Nếu không, đặt attack_flag là 0 (không có tấn công).

❖ **|table _time, srcip, dstip, total_sessions, total_traffic,packet_race, attack_flag**

- Hiển thị các cột: _time, srcip, dstip, total_sessions, total_traffic, packet_race,và attack_flag.



Hình3.47: Kết quả sau khi sử dụng thuật toán Logistic Regression.

Chúng em quyết định sử dụng thuật toán hồi quy Logistic(Logistic Regression) để phát hiện các lưu lượng Dos .Trong truy vấn có sử dụng attack_flag để biểu diễn bất thường hay không bất thường. Vì vậy hồi quy logistic dựa trên xác suất, dễ dàng đưa ra quyết định "bất thường" hay "không bất thường" bằng cách thiết lập một ngưỡng xác suất.

- Field to predict : attack_flag
- Fields to use for predicting: total_sessions, total_traffic,packet_race
- Split for training / test: **80 / 20**
  - **80% huấn luyện**: Đảm bảo mô hình có đủ dữ liệu để học và tìm ra mối quan hệ giữa các đặc trưng.
  - **20% kiểm tra**: Đảm bảo một phần dữ liệu không được mô hình "nhìn thấy" trong quá trình huấn luyện, dùng để đánh giá hiệu suất mô hình trên dữ liệu mới.

Splunk đã huấn luyện logistic regression và học được trọng số **w=[0.2,0.5,0.3]** và hệ số chặn: **b=-0.1**. Đây là các trọng số đã được huấn luyện mô hình cho các đặc trưng (features) trong dữ liệu.

❖ Tính toán xác suất với sigmoid:

- Các giá trị được lấy từ log ta có:
  - total_traffic=477986
  - packet_rate=9.4
  - total_sessions=515

- Giá trị z: z=(0.2×**477986)** +(0.5×**9.4)**+(0.3×**515)**-0.1 = 95756.3
- Xác suất bất thường: f(z) = $\frac{1}{1+e^{-95756.3}}$ ≈1

❖ f(z) = 1. Suy ra Splunk phân loại sự kiện này là bất thường (Vì f(z)>0.5f(z) > 0.5f(z)>0.5, mô hình sẽ phân loại đây là một tấn công (DoS attack)).

CHƯƠNG 4: KẾT LUẬN

Hệ thống quản lý Log và sự kiện tập trung SIEM (Security Information and Event Management) đóng vai trò quan trọng trong việc bảo vệ các tổ chức khỏi các mối đe dọa an ninh mạng ngày càng tinh vi. Trong quá trình tìm hiểu, nghiên cứu và triển khai,chúng em đã có những kiến thức hữu ích về các cơ chế hoạt động của hệ thống SIEM, từ việc thu thập và phân tích log đến việc cảnh báo và phát hiện các mối đe dọa an ninh mạng. Việc thu thập log từ nhiều thiết bị như Fireall,Webserver,ADserver về một hệ thống đã chứng tỏ được hiệu quả trong việc thu thập và phân tích mạnh mẽ của hệ thống Splunk.

Việc lựa chọn một hệ thống SIEM phù hợp cần dựa trên nhiều tiêu chí như khả năng mở rộng, dễ sử dụng, tích hợp với các hệ thống hiện có, tính năng tìm kiếm và phân tích, chi phí, và hỗ trợ từ nhà cung cấp. Các hệ thống như Splunk, AlienVault và OSSIM đều có những ưu điểm riêng cũng như các mặt hạn chế , việc chọn lựa hệ thống nào sẽ phụ thuộc vào nhu cầu cụ thể và khả năng tài chính của các tổ chức và doanh nhiệp.

Tuy nhiên, trong quá trình triển khai hệ thống SIEM, chúng em cũng gặp phải một số khó khăn trong việc thu thập và chuẩn hóa dữ liệu.Chuẩn hóa dữ liệu từ nhiều nguồn thiết bị khác nhau, mỗi thiết bị có cấu trúc log khác nhau. Điều này yêu cầu các câu truy vấn SPL trong Splunk phải luôn được điều chỉnh để đồng nhất hóa dữ liệu, giúp hệ thống phân tích dễ dàng hơn . Hơn nữa, việc xử lý các nguồn dữ liệu có sự khác biệt về cấu trúc và tần suất thu thập đòi hỏi một sự tinh chỉnh và tối ưu hóa liên tục trong quá trình triển khai. Để mở rộng mô hình và nâng cao khả năng phát hiện và ngăn chặn tấn công, ngoài việc triển khai hệ thống SIEM như Splunk, việc tích hợp các giải pháp IDS và IPS sẽ cung cấp nguồn dữ liệu quan trọng từ quá trình giám sát và phân tích mạng hỗ trợ hệ thống SIEM phân tích dữ liệu và đưa ra cảnh báo nhanh chóng, giảm thiểu tác hại từ các cuộc tấn công .

Tóm lại, SIEM là phần quan trọng trong chiến lược bảo mật của bất kỳ tổ chức nào, giúp bảo vệ dữ liệu và tài sản quan trọng mà còn nâng cao khả năng phản ứng và phục hồi trước các mối đe dọa bảo mật hiện đại. Các tổ chức và doanh nghiệp có thể tự tin hơn trong việc giám sát, phân tích và đảm bảo an toàn cho hệ thống dữ liệu khi sở hữu một hệ thống SIEM mạnh mẽ.

TÀI LIỆU THAM KHẢO

[1] OWASP Top 10 – 2021 .[online]

From: <https://owasp.org/Top10/ >

[2] Chrissy Kidd.SIEM: Security Information & Event Management Explained(October 12, 2023)[online]

From:<https://www.splunk.com/en_us/blog/learn/siem-security-information-event-management.html>

[3] Developer Guide for Splunk Cloud Platform and Splunk Enterprise.[online]

From:< https://dev.splunk.com/enterprise/docs/welcome/>

[4]IBMQRadar:QradarUserGuide .[online]

From:< https://www.ibm.com/docs/en/SS42VS_7.4/pdf/b_qradar_users_guide.pdf >

[5] Juan Manuel Lorenzo. AlienVault Users Manual(2010-2011)[online]

From:< AlienVault_Users_Manual_1.0.pdf>

[6] Splunk Machine Learning Toolkit Showcase.[online]

From:<https://docs.splunk.com/Documentation/MLApp/5.5.0/User/Showcaseexamples>

PHỤ LỤC

STT	Kỹ tự chữ viết tắt	Nội dung
1	GDPR	General Data Protection Regulation là Quy định chung về bảo vệ dữ liệu cá nhân và quyền riêng tư của công dân EU được ban hành bởi Liên minh Châu Âu (EU).
2	Nghị định 85/2016/NĐ-CP	Nghị định của Chính phủ Việt Nam được ban hành vào ngày 1 tháng 7 năm 2016, quy định chi tiết về bảo đảm an toàn hệ thống thông tin theo cấp độ
3	ISO 27001	Tiêu chuẩn quốc tế về quản lý an ninh thông tin, được thiết kế để bảo vệ thông tin của tổ chức khỏi các mối đe dọa và đảm bảo sự an toàn, tính bảo mật và toàn vẹn của thông tin.
4	PCI-DSS	Một bộ tiêu chuẩn bảo mật được phát triển để bảo vệ thông tin thanh toán thẻ (credit/debit card) nhằm giảm nguy cơ gian lận và lộ lọt thông tin thẻ