HỆ THỐNG GIÁM SÁT AN NINH MẠNG SỬ DỤNG BỘ CÔNG CỤ ELK VÀ IDS

Võ Tá Hoàng, Phạm Mạnh Tùng
Trường Đại học Thủy lợi, email: hoangvota@tlu.edu.vn

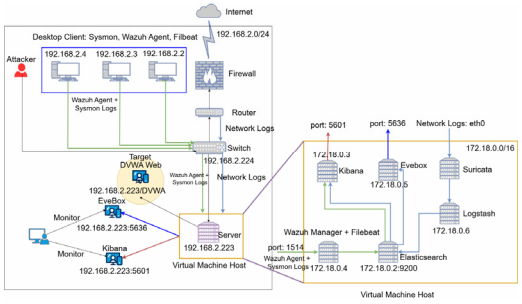
1. GIỚI THIỆU CHUNG

Hiện nay, sự gia tăng nhanh chóng của các dịch vụ dựa trên Cloud và các ứng dụng IoT khiến nhu cầu về tính khả dụng và khả năng tiếp cận tốt hơn giữa các nút internet cũng tăng lên. Đồng thời, điều này cũng làm tăng bề mặt tấn công của hệ thống, do các lỗ hổng của cơ sở hạ tầng hoặc phần mềm tạo ra. Để ngăn chặn những tình huống như vậy, cần có thiết bị phần cứng và công cụ phần mềm có khả năng phát hiện các mối đe dọa trong thời gian thực [1]. Vì vậy, bài báo này nghiên cứu các công cụ bảo mật phổ biến và cách chúng tích hợp với nhau để xây dựng hệ thống thu thập, phân tích, xử lý và hiển thị log phục vụ cho giám sát an ninh mạng.

2. KIẾN TRÚC HỆ THỐNG

Nghiên cứu này sử dụng bộ công cụ ELK và IDS để xây dựng hệ thống giám sát an ninh mạng. Elasticsearch là một công cụ phân tích và tìm kiếm phân tán được xây dựng trên Apache Lucene. Công cụ này được sử dụng để phân tích nhật ký, tìm kiếm văn bản, thông tin bảo mật, thông tin phân tích nghiệp vụ. Kibana là ứng dụng web để tìm kiếm và hiển thị dữ liệu từ Elasticsearch. Logstash là công cụ để thu thập, xử lý, chuẩn hóa và hợp nhất các thông báo sự kiện và nhật ký từ nhiều nguồn khác nhau [2]. EveBox cũng kết nối với Elasticsearch và cung cấp giao diện để phân tích và quản lý các sự kiện an ninh mạng. EveBox giúp kiểm tra các sự kiện phát hiện xâm nhập và các log mạng từ Suricata.

Hệ thống bao gồm một mạng người dùng nội bộ và một mạng máy ảo. Mạng máy ảo được cài đặt trên máy chủ và chứa các nút thu thập và xử lý dữ liệu được xây dựng bởi công cụ Elasticsearch, Logstash, Kibana, Suricata, Sysmon, Filebeat và Wazuh. Mô tả chi tiết của hệ thống như Hình 1.



Hình 1. *Kiến trúc hệ thống giám sát*

Mạng người dùng nội bộ gồm: Các máy tính, tường lửa, switch và máy chủ. Máy chủ chứa các máy ảo và dịch vụ giám sát. Trong đó, DVWA (Damn Vulnerable Web Application) là ứng dụng web chứa các lỗ hổng bảo mật để thử nghiệm tấn công; Evebox Web hiển thị các cảnh báo từ Suricata giúp quản trị viên có thể theo dõi các cảnh báo mạng một các trực quan hơn; Kibana Web hiển thị logs, cảnh báo, dữ liệu phân tích từ Elasticsearch. Đồng thời, tích hợp các dashboard để quản trị viên có thể đánh giá sự cố, theo dõi hoạt động của hệ thống một cách thuận tiện nhất.

Mạng máy ảo được sử dụng cho nút chứa công cụ ELK và IDS để thu thập và xử lý dữ liệu. Mô hình thu thập log trong hệ thống này có 2 luồng:

- **Luồng 1:** Suricata → Logstash → Elasticsearch → Evebox & Kibana

- Luồng 2: Wazuh Agent → Wazuh Manager + Filebeat → Elasticsearch → Kibana

Trong đó, đầu vào và đầu ra của mỗi công cụ thành phần như Bảng 1 dưới đây:

Bảng 1. Mô tả dữ liệu của các công cụ

Công cụ	Dữ liệu đầu vào	Dữ liệu đầu ra
Suricata	Lưu lượng mạng dạng IP packet	Log sự kiện dạng JSON
Logstash	Log sự kiện dạng JSON	Log đã được phân tích dạng JSON
Elasticsearch	Log đã được phân tích dạng JSON	Log được lập chỉ mục dạng JSON
Evebox	Log được lập chỉ mục dạng JSON	Hiển thị bảng dữ liệu an ninh mạng cơ bản
Kibana	Log được lập chỉ mục dạng JSON	Hiển thị bảng, biểu đồ nâng cao
Wazuh agent	Log file của hệ thống trên host	Dữ liệu log được cấu trúc lại dưới dạng JSON
Wazuh manager + File Beat	Log từ Wazuh agent dạng JSON	Log được chuẩn hóa dạng JSON

3. KỊCH BẢN TẤN CÔNG VÀ KẾT QUẢ

3.1. Mục tiêu

Kịch bản tấn công này nhằm kiểm tra và đánh giá khả năng giám sát, phát hiện và phản ứng của hệ thống đối với cuộc tấn công File Upload trên ứng dụng web DVWA.

3.2. Kịch bản và phương thức tấn công

Trang web DVWA chạy trên máy chủ: 192.168.2.223/DVWA.

Trang web này tồn tại các lỗ hổng: Brute Force, Command Injection, CSRF, File Inclusion, File Upload, Insecure CAPTCHA, SQL Injection, SQL Injection (Blind), Weak Sessions IDs, XSS (DOM), XSS (Reflected), XSS (Stored), CSP Bypass, JavaScript, Authorisation Bypass, Open HTTP Redirect.

Kẻ tấn công: là một tác nhân nội bộ và thực hiện cuộc tấn công từ một máy tính nội bộ và tấn công vào lỗ hổng trên trang web.

Bước 1: Chuẩn bị môi trường

- Truy cập DVWA

Mở trình duyệt web và truy cập vào trang đăng nhập của DVWA: http://192.168.2.223/DVWA với 1 tài khoản.

Bước 2: Tấn công File Upload

- Tìm kiếm lỗ hổng File Upload: Truy cập vào phần "File Upload" của DVWA. Chuẩn bị một tệp PHP độc hại (ví dụ: shell.php) với nội dung đơn giản như <?php system($_GET['cmd']); ?>. Upload tệp shell.php lên DVWA.

- Khai thác lỗ hổng File Upload: Truy cập tệp đã tải lên qua đường dẫn tương ứng trên DVWA, ví dụ: http://192.168.2.223/DVWA/hackable/uploads/shell.php?cmd=ls.

- Ghi nhận log: Suricata phát hiện và ghi nhận các mẫu lưu lượng bất thường. Logstash xử lý và chuyển log tới Elasticsearch. Kibana và EveBox hiển thị log và cung cấp phân tích.

Bước 3: Phân tích log và cảnh báo

- Phân tích trên Kibana: Mở Kibana tại http://192.168.2.223:5601 và phân tích log hiển thị trên dashboard. Tùy thuộc loại cảnh báo trên dashboard, người giám sát sẽ tiến hành phân tích cụ thể từng loại log khi có cảnh báo. Trong bài báo này, ví dụ về phân tích log được chỉ ra trong kết quả của Hình 2 và Hình 3.

Bước 4: Báo cáo và phản hồi

- Tạo báo cáo tấn công: Tạo báo cáo chi tiết về các cuộc tấn công đã thực hiện, bao gồm phương pháp, các log ghi nhận, và phân tích.

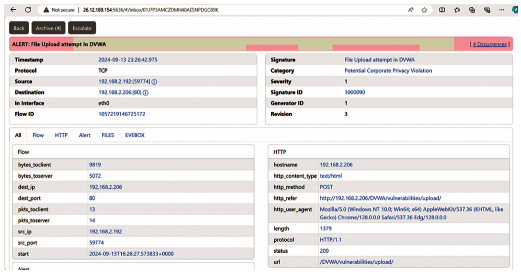
- Phản hồi và cải thiện: Dựa trên báo cáo, xác định các điểm yếu trong hệ thống. Đề xuất các biện pháp cải thiện và bảo vệ hệ thống khỏi các cuộc tấn công tương tự trong tương lai. Tìm kiếm các sự kiện bất thường và các yêu cầu chứa các chuỗi độc hại.

- Phân tích trên EveBox: Mở EveBox tại http://192.168.2.223:5636. Xem các sự kiện an ninh mạng được ghi nhận từ Suricata. Phân tích các sự kiện phát hiện xâm nhập liên quan đến các cuộc tấn công.

3.3. Kết quả

Hình 2. *Kết quả hiển thị cảnh báo tệp độ hại được tải lên web server*

Hình 2 cho thấy hệ thống hoạt động chính xác khi đã phát hiện và cảnh báo kịp thời sự kiện có tệp độc hại được tải lên web server khi chúng tôi giả lập một cuộc tấn công mã độc vào web server.



Hình 3. *Chi tiết của cảnh báo bị tấn công*

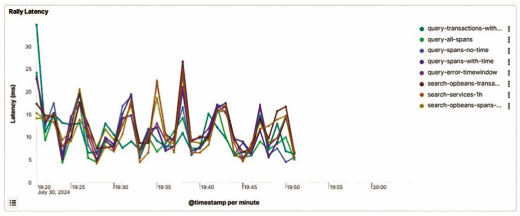
Hình 3 chỉ rõ thông tin chi tiết của cảnh báo bị tấn công mã độc như địa chỉ IP nguồn/đích, cổng nguồn/đích, đường dẫn, thời gian, chi tiết HTTP request tải lên tệp độc hại, chi tiết về luồng,...

Ngoài ra, nghiên cứu này đã sử dụng công cụ Rally để đo hiệu năng của hệ thống. Công cụ này tạo giả lập tải cho hệ thống và đo các thông số đánh giá hiệu năng hệ thống.

Bảng 2. Hiệu năng của hệ thống

Operations	100th Percentile (ms)	90th Percentile (ms)	50th Percentile (ms)
query-transactions-with-...	29.549	10.49	7.27
query-all-spans	26.681	11.035	7.151
query-spans-no-time	24.489	9.733	4.761
query-spans-with-time	19.858	8.035	4.245
search-opbeans-spans-2...	15.687	10.901	7.352
query-error-timewindow	14.988	11.215	7.381
search-opbeans-transact...	10.981	9.879	6.731
search-services-1h	10.886	8.72	5.594

Bảng 2 là kết quả đo đạc hiệu năng của hệ thống với các loại yêu cầu và điều kiện tải khác nhau. Trong đó, 100th Percentile là thời gian phản hồi hoặc độ trễ cao nhất trong toàn bộ các yêu cầu. 90th Percentile và 50th Percentile lần lượt là thời gian phản hồi hoặc độ trễ mà 90% và 50% số yêu cầu có thời gian phản hồi nhanh hơn.



Hình 4. *Độ trễ của hệ thống*

Kết quả độ trễ hệ thống tại Hình 4 biến động theo thời gian bởi vì khi kiểm tra hiệu năng, Rally giả lập các yêu cầu theo nhiều loại phân bố khác nhau để mô phỏng các tình huống trong thực tế, và các phân bố này sẽ giúp cho kiểm tra hiệu năng của hệ thống với các điều kiện tải khác nhau, từ đó cung cấp cái nhìn toàn diện về hiệu năng của hệ thống trong các tình huống thực tế. Các phân bố được sử dụng như phân bố đều, phân bố Poisson, phân bố mũ, phân bố bậc thang [3].

4. KẾT LUẬN

Chúng tôi đã xây dựng được một hệ thống quản lý, giám sát các sự kiện và thông tin an ninh mạng nhằm phục vụ cho việc cảnh báo sớm các sự cố hoặc hành vi vi phạm an toàn thông tin đối với hệ thống.

5. TÀI LIỆU THAM KHẢO

[1] Negoita, Ovidiu, et al., 2020, "Enhanced security using elasticsearch and machine learning." Intelligent Computing: Proceedings of the 2020 Computing Conference, Volume 3. Springer International Publishing.

[2] Ahmed, Farrukh, et al., 2020, "Centralized log management using elasticsearch, logstash and kibana." 2020 International Conference on Information Science and Communication Technology (ICISCT). IEEE.

[3] Haugerud, Hårek, et al., 2022, "Tuning of elasticsearch configuration: parameter optimization through simultaneous perturbation stochastic approximation." Frontiers in big Data 5.