# Báo cáo cuối kì

Nguyễn Bình Thục Trâm[20520815], Nguyễn Bùi Kim Ngân[20520648], and Võ Anh Kiệt[20520605]

UIT, VNU-HCM, Ho Chi Minh city, Vietnam
20520815@gm.uit.edu.vn, 20520648@gm.uit.edu.vn, 20520605@gm.uit.edu.vn

**Tóm tắt nội dung** Báo cáo này tập trung vào nghiên cứu và đánh giá hiệu suất của hai Hệ thống Phát hiện và Ngăn chặn Xâm nhập (IDPS) mã nguồn mở, Snort và Suricata, trong việc phát hiện và ngăn chặn các tấn công mạng thông thường như ping, hping3 (syn flood), hydra dò mật khẩu, telnet, và nmap. Nhóm đã xây dựng các rule tương ứng để kiểm tra tính năng của cả Snort và Suricata, đồng thời triển khai một Mô hình Machine Learning (ML-based IDPS) để phát hiện mã độc. Phương pháp này nhằm đánh giá sự hiệu quả và tính linh hoạt của các công cụ IDPS cũng như khả năng phát triển của hệ thống dựa trên học máy.

**Keywords:** IDS · IPS · Machine learning

## 1 Giới thiệu chung

### 1.1 Tổng quan về IDPS

Hệ thống tìm kiếm, phát hiện và ngăn ngừa xâm nhập (IDPS) là hệ thống bảo mật mạng được sử dụng tập trung chủ yếu vào việc xác định các sự cố có thể xảy ra, ghi nhật ký thông tin về chúng, cố gắng ngăn chặn và báo cáo chúng cho quản trị viên bảo mật. IDPS được chia làm 3 loại chính dựa trên data source: **Network-based IDPS (NIDPS)** giám sát lưu lượng mạng trên các phân đoạn hoặc thiết bị mạng cụ thể, phân tích hoạt động giao thức mạng và ứng dụng để xác định hoạt động đáng ngờ; **Host-based IDPS (HIDPS)** giám sát các đặc điểm của một host và các sự kiện xảy ra trong host đó để phát hiện hoạt động đáng ngờ (lưu lượng truy cập mạng (chỉ dành cho host đó), nhật ký hệ thống, tiến trình đang chạy, hoạt động ứng dụng, truy cập và sửa đổi tệp, và những thay đổi cấu hình của hệ thống, ứng dụng); **Hybrid IDS** được phát triển dựa trên dữ liệu được cung cấp bởi các sự kiện trên host và các phân đoạn mạng, đồng thời kết hợp những chức năng của cả 2 loại NIDPS và HIDPS.
Về phương pháp phát hiện xâm nhập gồm có:

- Signature-based (or misuse, knowledge-based): So sánh dữ liệu với các mẫu đã biết. Có ưu điểm là độ chính xác phát hiện cao đối với các mối đe dọa đã biết, tỷ lệ cảnh báo sai thấp. Tuy nhiên nhược điểm là độ chính xác phụ thuộc vào cơ sở dữ liệu signature, không thể phát hiện những bất thường chưa xác định hoặc những biến thể của cuộc tấn công đã biết.

- Anomaly-Based (or profile-based): Tạo và so sánh với baseline profile để xác định hành vi bất thường. Ưu điểm là có thể phát hiện được bất thường đã và chưa biết cũng như phát hiện được những tấn công mới. Hạn chế là tỷ lệ cảnh báo sai và bỏ sót cao, tiêu tốn nhiều tài nguyên.
- Specification-based: Giám sát các chương trình đang thực thi để phát hiện bất thường từ các thông số kỹ thuật hợp lệ tương ứng.
- Hybid: Kết hợp các loại trên. Ưu điểm mang lại là chống được những tấn công có thay đổi tinh vi, tổng hợp những lợi ích của mỗi loại và khắc phục được nhiều nhược điểm.

IDPS là một thành phần quan trọng trong hệ thống bảo mật mạng, giúp bảo vệ mạng máy tính khỏi các cuộc tấn công mạng, từ đó bảo vệ dữ liệu và tài sản của tổ chức.

### 1.2 Hướng nghiên cứu của nhóm

Phần này sẽ giới thiệu về những công cụ IDPS mà nhóm đã tiến hành thử nghiệm thử nghiệm, bao gồm:

- Snort: Một trong những công cụ IDPS lâu đời và phổ biến nhất hiện nay
- Suricata: Công cụ miễn phí, có mã nguồn mở và chạy được trên nhiều nền tảng
- ML-based IDPS: Thực hiện cài đặt các phương pháp học máy vào trong hệ thống IDPS để phát hiện mã độc

Nhóm chạy thử nghiệm để kiểm tra một số tính năng cũng như độ hiệu quả của chúng và đưa ra đánh giá chung cho cả 3 công cụ trên.

## 2 Tổng quan về các công cụ nhóm sử dụng

### 2.1 Snort

**Giới thiệu** Snort là hệ thống Phát hiện và Ngăn chặn Xâm nhập (IDPS), mã nguồn mở, có ngôn ngữ lập trình C, được phát triển bởi Martin Roesch, người sáng lập cựu CTO của Sourcefire vào năm 1991 và hiện được phát triển bởi Cisco. Là một công cụ mạnh mẽ hỗ trợ nhiều nền tảng Linux, Windows, và macOS, có thể phát hiện nhiều loại cuộc tấn công mạng.

Snort có thể được cấu hình ở ba mode chính gồm:
1. Sniffer Mode: đọc các gói mạng và hiển thị chúng trên bảng điều khiển.
2. Packet Logger Mode: ghi lại các packet vào đĩa.
3. Network Intrusion Detection System Mode: giám sát lưu lượng mạng và phân tích nó theo bộ quy tắc do user định nghĩa. Sau đó thực hiện một hành động cụ thể dựa trên những gì đã được xác định.

**Tính năng** Một số tính năng chính của Snort bao gồm:

- Phát hiện xâm nhập mạng (IDS): Snort có thể phân tích lưu lượng mạng thời gian thực để phát hiện các hoạt động đáng ngờ và các cuộc tấn công như scans, exploits, DDoS, etc.
- Ngăn chặn xâm nhập mạng (IPS): Snort có thể được cấu hình để chặn và ngăn chặn các gói tin liên quan đến tấn công. Điều này giúp ngăn ngừa các cuộc tấn công thành công.
- Phân tích giao thức và phát hiện bất thường: Snort có khả năng phân tích hàng trăm loại giao thức mạng khác nhau để phát hiện các hoạt động bất thường như vi phạm chính sách, lạm dụng giao thức, etc.
- Báo cáo và ghi log: Snort ghi lại chi tiết các sự kiện phát hiện được và có thể tạo các báo cáo dễ hiểu để phân tích.
- Cảnh báo thời gian thực: Snort có thể được cấu hình để cảnh báo ngay lập tức khi phát hiện đột nhập hoặc tấn công.

### 2.2 Suricata

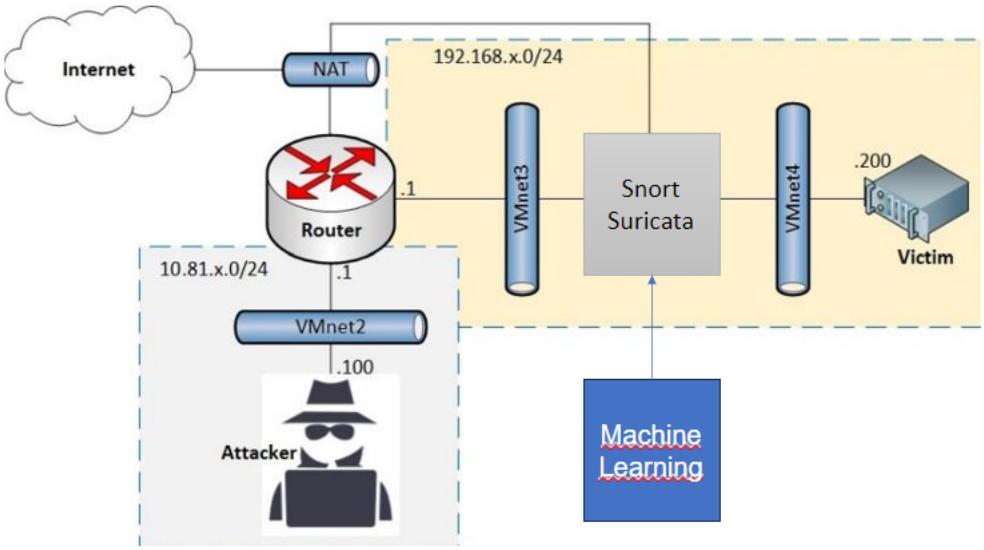
**Giới thiệu** Suricata là một dự án được khởi động vào năm 2008 bởi nhóm phát triển tại Open Information Security Foundation (OISF) với mục tiêu tạo ra một giải pháp phát hiện xâm nhập và ngăn chặn xâm nhập mới, mở rộng, và linh hoạt. Đến tháng 12 năm 2010, phiên bản Suricata 1.0.0 được phát hành đánh dấu sự ra đời của một hệ thống phát hiện xâm nhập (IDS) và hệ thống ngăn chặn xâm nhập (IPS) mã nguồn mở có khả năng giám sát và bảo vệ các hệ thống mạng mạng khỏi các mối đe dọa mạng. Nó được phát triển chủ yếu để hiệu quả trong môi trường mạng cao băng thông và có thể tích hợp với các hệ thống khác như Snort.

**Tính năng** Suricata cho thấy rằng mình là một hệ thống phát hiện và ngăn ngừa xâm nhập mạng hiệu quả với các tính năng nổi trội:

- Suricata có khả năng bắt gói tin ở cả chế độ không lưu trạng thái (stateless) và lưu trạng thái (stateful). Cung cấp khả năng kiểm tra mỗi gói tin độc lập hoặc dựa trên thông tin kết nối trước đó.
- Suricata hỗ trợ nhiều giao thức khác nhau như ICMP, DNS, HTTP, TLS, FTP,...
- Suricata cung cấp nhiều tùy chọn cấu hình và quy tắc để tùy chỉnh theo nhu cầu cụ thể của hệ thống. Điều này bao gồm cả khả năng kích thích hoặc giả mạo các cuộc tấn công để ghi log và phân tích.
- Suricata có thể xử lý đa luồng mạnh mẽ, tăng khả năng mở rộng và hiệu suất của hệ thống trong môi trường mạng phức tạp.
- Suricata có hệ thống ghi log đủ khả năng cung cấp các tệp tin log chi tiết về các sự kiện xâm nhập và các hoạt động của mạng, giúp nhận biết, phản ứng nhanh chóng, thực hiện truy vết đối với các mối đe dọa.

# 3 Phương pháp nghiên cứu

## 3.1 Mô hình tổng quan




**Hình 1.** Mô hình tổng quan

## 3.2 Tools IDPS: Snort, Suricata

Đây là phương pháp nghiên cứu IDS/IPS qua công cụ Snort và Suricata:

- Cài đặt Snort/Suricata trên hệ thống (cài đặt các gói liên quan, cấu hình các file cần thiết)
- Thực hiện chạy thử trước khi cài rules
- Viết rules để cảnh báo khi có lưu lượng bất thường, chặn gói tin độc hại.
- Khởi động dịch vụ Snort/Suricata
- Thử nghiệm các kịch bản tấn công đến để xem Snort/Suricata có phát hiện, cảnh báo và ngăn chặn được các cuộc tấn công hay không.

Cài đặt và apply rules:

- Tìm, thu thập các rulesets cho Snort/Suricata trên mạng
- Nghiên cứu và viết thêm một số rules đơn giản cho riêng mình
- Áp dụng các rulesets đó vào Snort/Suricata
- Load lại các rule vừa thêm vào

Kiểm tra kết quả sau khi áp dụng rules:

- Chuẩn bị kịch bản và môi trường để thực hiện các cuộc tấn công.
- Thực hiện tấn công đến các máy đã cài đặt Snort/Suricata.
- Xem trong real-time logs của Snort/Suricata để kiểm tra khả năng real-time alert các lưu lượng được cho là độc hại dựa trên rules đã thêm vào.
- Kiểm tra trên máy tấn công, máy cài đặt Snort/Suricata để xem các dấu hiệu cho thấy rằng cuộc tấn công đã bị ngăn chặn.
- Như vậy là đã có thể nghiên cứu hoạt động của IDS/IPS thông qua việc cài đặt, cấu hình và áp dụng rules cho Snort/Suricata.

### 3.3 ML-based IDPS: Malware detection

Đây là phương pháp nghiên cứu sử dụng Machine Learning để phát hiện mã độc trong IDS/IPS:

**Thu thập tập dữ liệu:**

- Truy cập Kaggle hoặc các nguồn dữ liệu khác để tải các dataset về mã độc
- Tập dữ liệu nên có đủ các thuộc tính cần thiết để huấn luyện mô hình, ví dụ: file signature, network traffic features, labels...

**Tiền xử lý dữ liệu:**

- Tách dữ liệu train/test
- Làm sạch dữ liệu như xử lý các giá trị thiếu, nhiễu, cân bằng tập dữ liệu...
- Trích chọn, biến đổi thuộc tính

**Xây dựng mô hình học máy:**

- Thuật toán: Logistic Regression, SVM, CNN, RNN, Random Forest...
- Huấn luyện mô hình trên tập train
- Đánh giá mô hình trên tập test

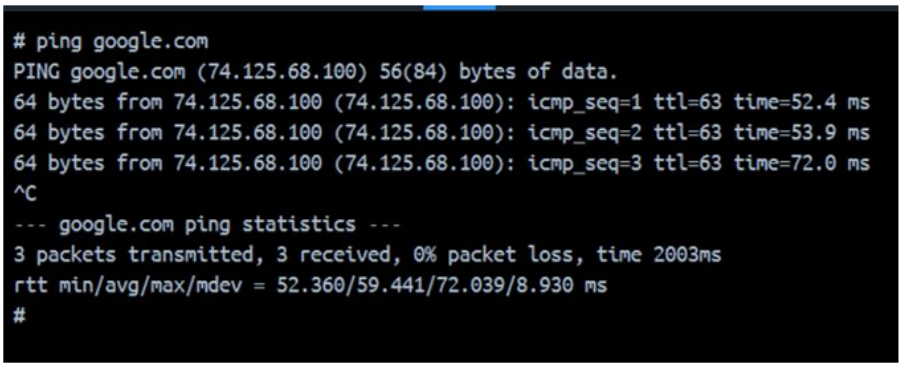
**Đánh giá và cải tiến mô hình**

- Đánh giá mô hình dựa trên các chỉ số precision, recall, f1-score, accuracy...
- Thử nghiệm nhiều mô hình khác nhau để tìm ra mô hình tốt nhất
- Tiến hành cải tiến dựa trên những giai đoạn trên nhằm nâng cao độ chính xác
- Như vậy đã có thể xây dựng được mô hình học máy phát hiện mã độc để áp dụng trong IDS/IPS

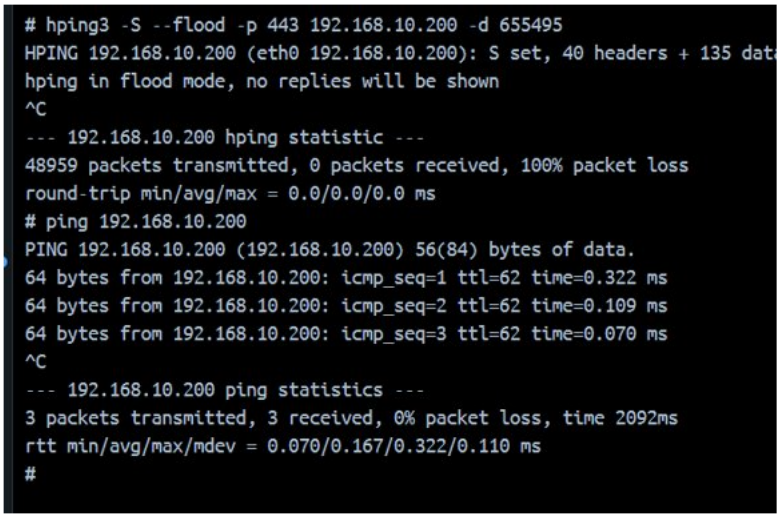
## 4 Kết quả thực nghiệm

### 4.1 Snort

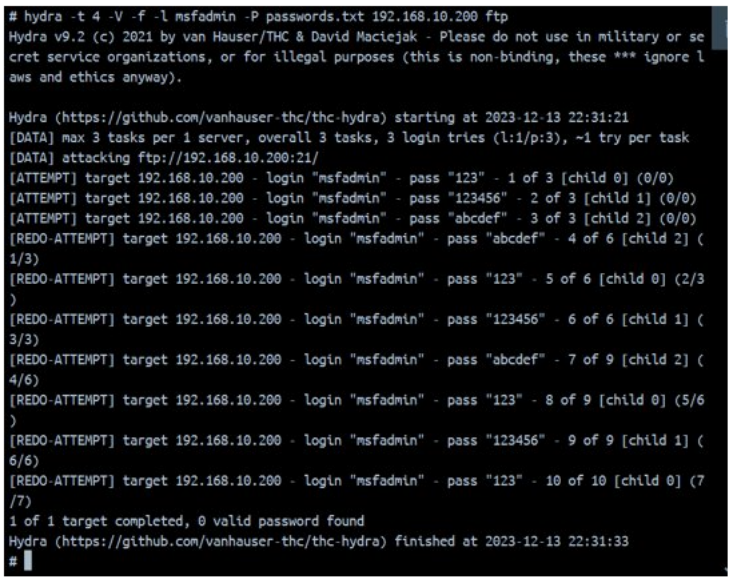
Trước khi cài rule, hình ảnh minh họa:



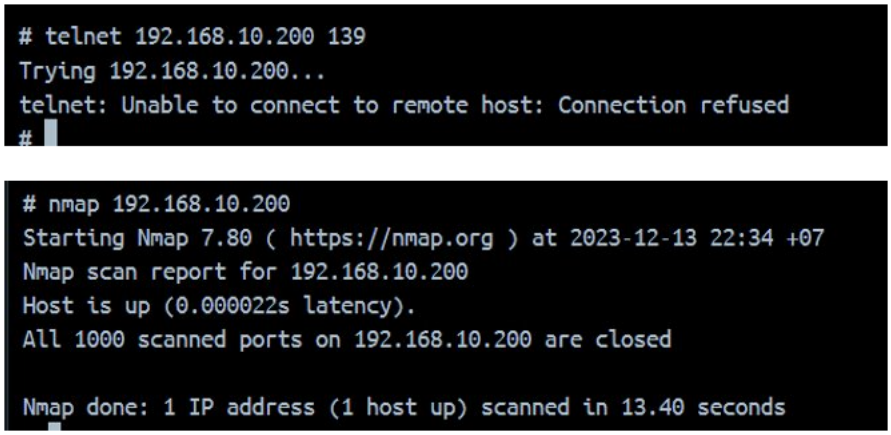

**Hình 2.** ping



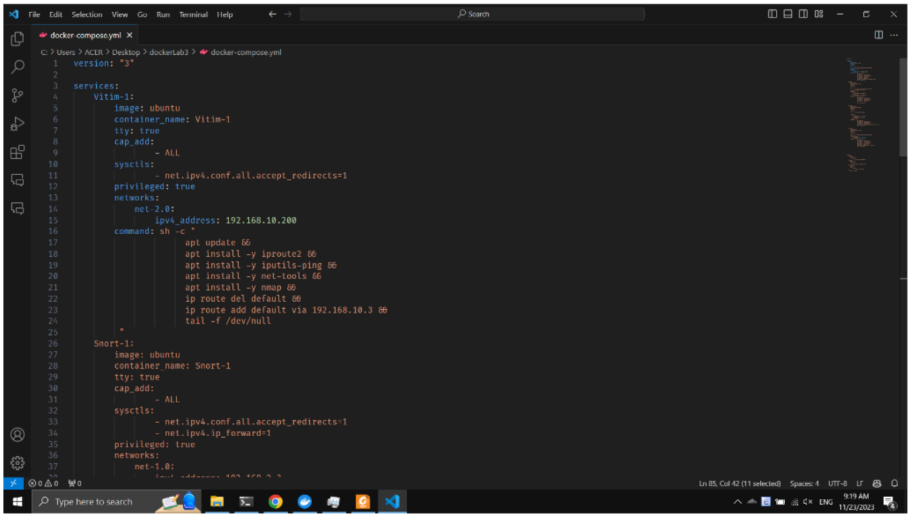

**Hình 3.** hping3



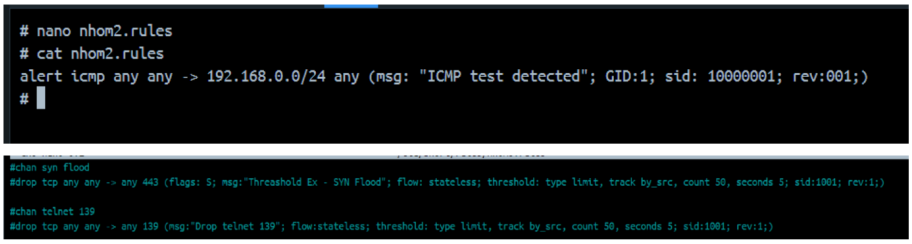

**Hình 4.** hydra



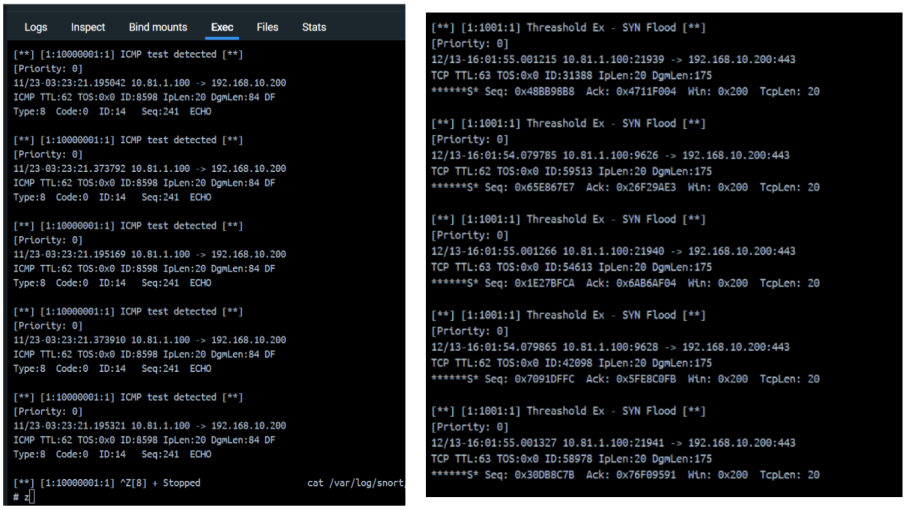

**Hình 5.** telnet, nmap



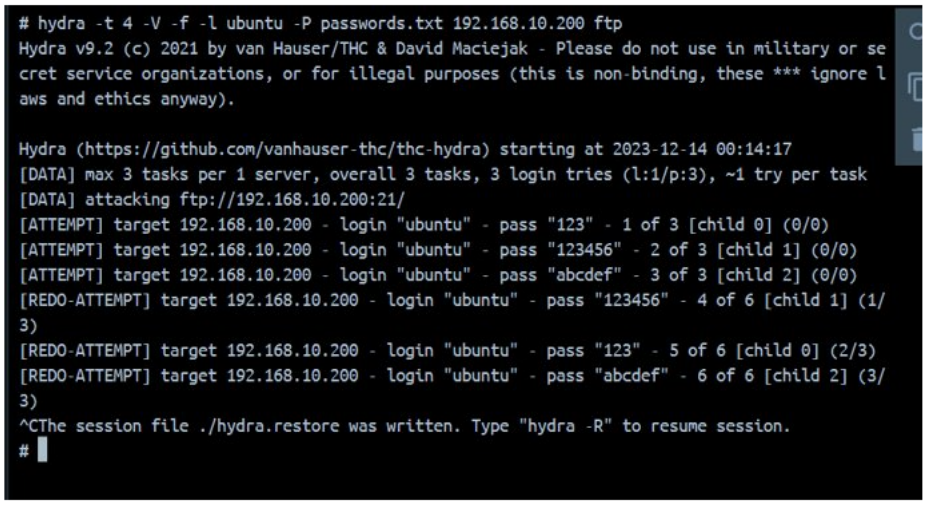

**Hình 6.** set up snort



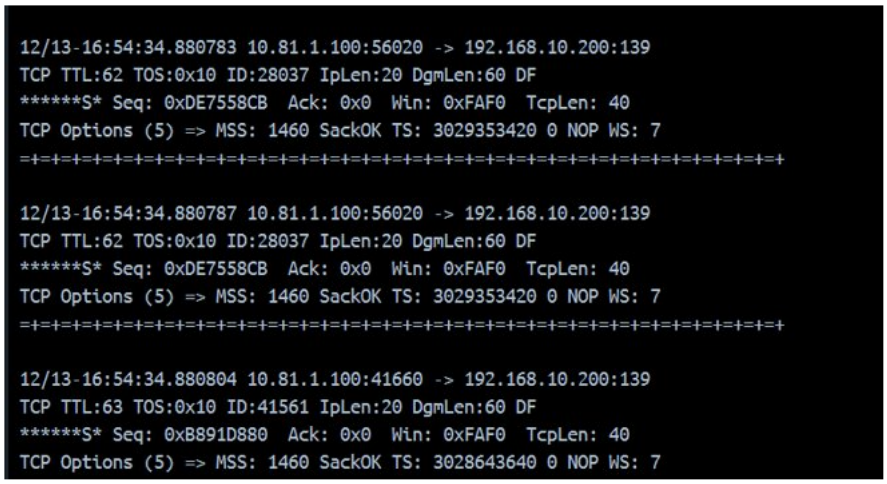

**Hình 7.** rule snort




**Hình 8.** Detect ping and hping3




**Hình 9.** Detect Hydra

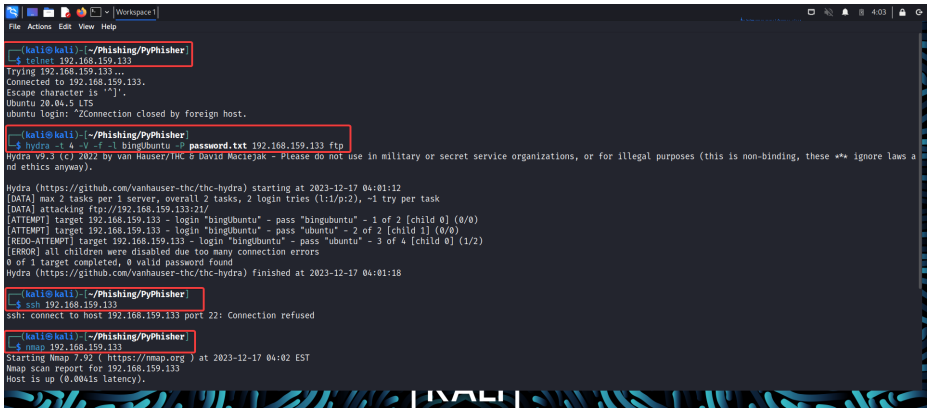



**Hình 10.** Detect telnet

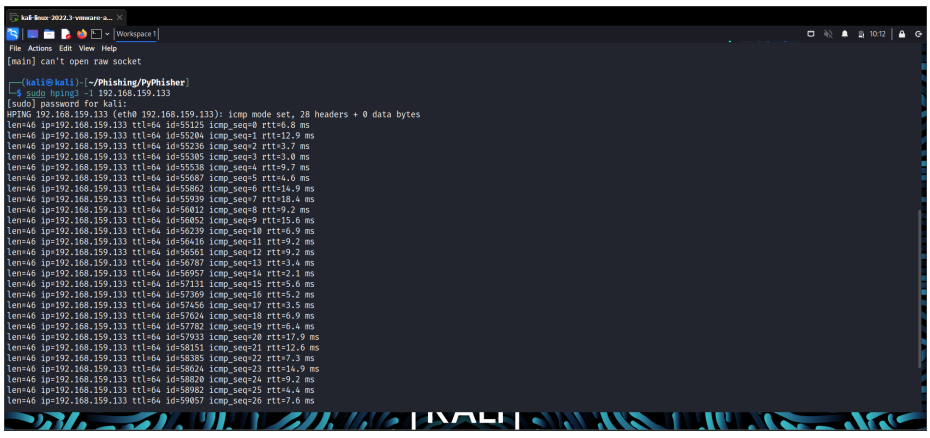
```
# nmap 192.168.200 -p 80-443
Starting Nmap 7.80 ( https://nmap.org ) at 2023-12-14 04:14 UTC
Note: Host seems down. If it is really up, but blocking our ping probes, try -Pn
Nmap done: 1 IP address (0 hosts up) scanned in 3.19 seconds
#
```

**Hình 11.** Detect nmap

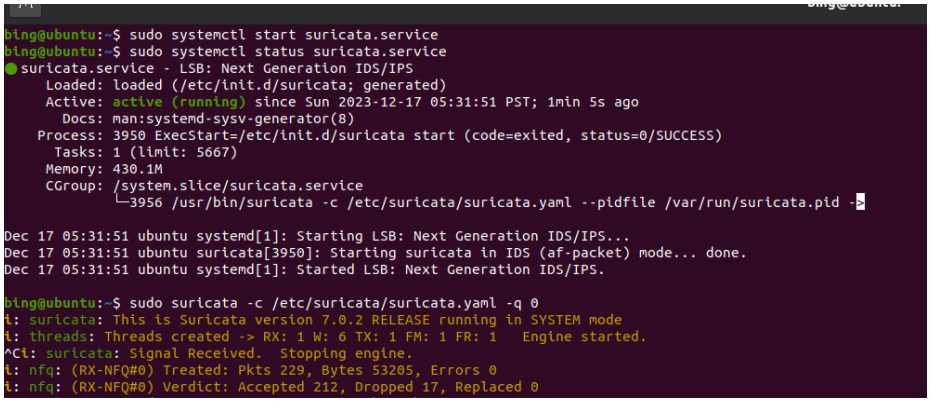
## 4.2 Suricata



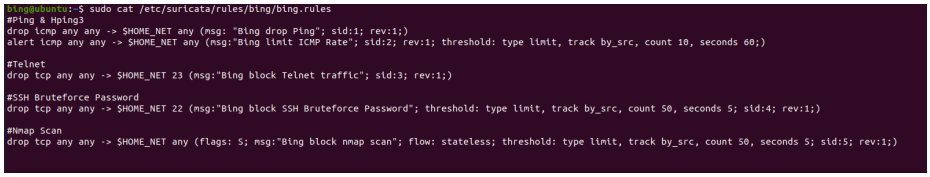

**Hình 12.** Trước khi tấn công, các hoạt động bình thường



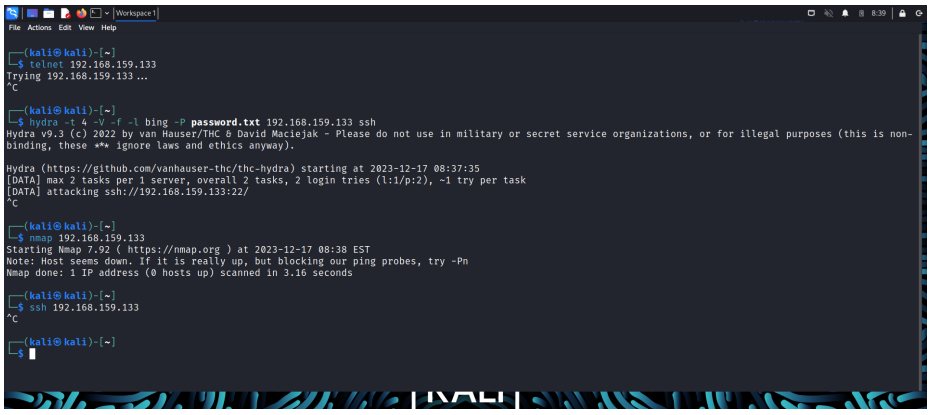

**Hình 13.** Trước khi tấn công, các hoạt động bình thường



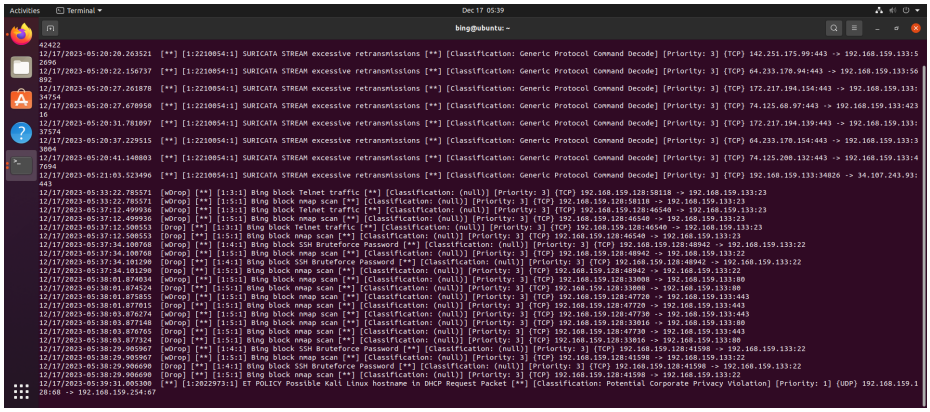

**Hình 14.** Cài đặt Suricata



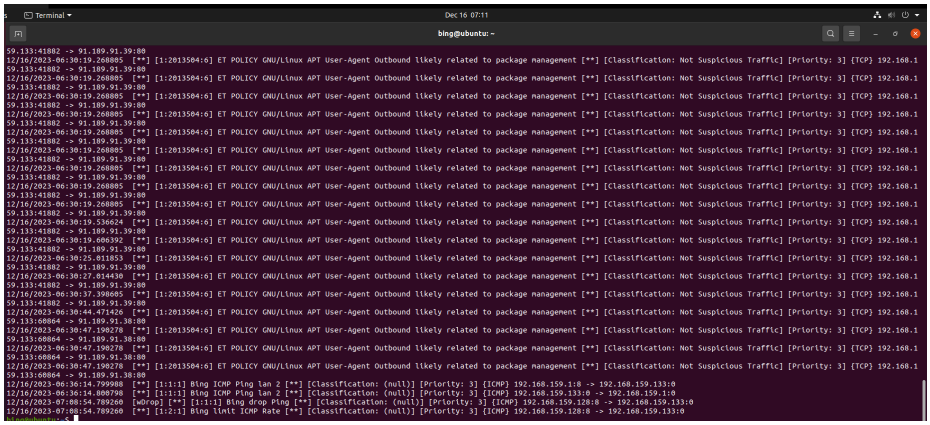

**Hình 15.** Thiết lập rule




**Hình 16.** Phát hiện tấn công



**Hình 17.** Log suricata



**Hình 18.** Log suricata

## 4.3 ML-based IDPS: Malware detection

**Bảng 1.** Kết quả phát hiện malware của ML-based IDPS với các model khác nhau

| Model | Accuraccy | Recall | Precision | F1Score | FAR |
|---|---|---|---|---|---|
| DecisionTree | 1.00 | 1.00 | 1.00 | 1.00 | 0.00 |
| KNN | 1.00 | 1.00 | 1.00 | 1.00 | 0.00 |
| LogisticRegression | 0.94 | 0.93 | 0.95 | 0.94 | 0.07 |
| RandomForest | 1.00 | 1.00 | 1.00 | 1.00 | 0.00 |
| SVM | 0.95 | 0.94 | 0.96 | 0.95 | 0.06 |
| XGBoost | 1.00 | 1.00 | 1.00 | 1.00 | 0.00 |
| CNN | 1.00 | 1.00 | 1.00 | 1.00 | 0.00 |
| FeedForwardNeuralNetwork | 1.00 | 1.00 | 1.00 | 1.00 | 0.00 |
| LSTM | 1.00 | 1.00 | 1.00 | 1.00 | 0.00 |

## 5 Kết luận

Trong quá trình thực hiện đồ án, nhóm đã tận dụng Snort và Suricata là 2 công cụ IDPS mã nguồn mở, phổ biến hiện nay để triển khai và đánh giá hiệu suất trong việc phát hiện và ngăn chặn các tấn công phổ biến. Việc xây dựng các rule cho các mô hình tấn công như ping, syn flood, dò mật khẩu, telnet, và nmap giúp nhóm kiểm tra và so sánh tính năng của cả hai công cụ. Đồng thời, việc triển khai thêm một ML-based IDPS giúp chúng tôi mở rộng khả năng phát hiện malware của IDPS. Kết quả thu được từ các thử nghiệm và đánh giá sẽ cung cấp cái nhìn chi tiết và toàn diện về tính năng và hiệu suất của từng công cụ, đồng thời xem xét tiềm năng và thách thức trong việc tích hợp mô hình học máy vào các hệ thống phát hiện và ngăn ngừa xâm nhập.