HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG
KHOA AN TOÀN THÔNG TIN

Báo cáo bài thực hành

Phát hiện và phản ứng sự cố tấn công quét thư mục web (web directory brute-force)

Sinh viên thực hiện:

B20DCAT002 – Hoàng Thu Cúc

Giảng viên hướng dẫn: TS.Nguyễn Ngọc Điệp

HÀ NỘI 12-2025

MỤC LỤC

DANH MỤC CÁC HÌNH VẼ

NỘI DUNG THỰC HÀNH

1.1 Lab 3: Phát hiện và phản ứng sự cố tấn công quét thư mục web (web directory brute-force)

1.1.1 Giới thiệu chung

Bài thực hành "Phát hiện và phản ứng sự cố tấn công quét thư mục web (web directory brute-force)" được xây dựng nhằm giúp sinh viên :

- Cấu hình Splunk để thu thập log web
- Theo dõi spike HTTP 404/403 và xác định IP tấn công
- Chặn tạm thời bằng ModSecurity

1.1.2 Mục đích

Sinh viên biết cách:

- Phát hiện và xử lý tấn công web directory brute-force bằng Splunk
- Phản ứng trước nguy cơ dò quét thư mục

1.1.3 Yêu cầu đối với sinh viên

- Nắm được kiến thức về HTTP, log web Apache và mã trạng thái 404/403.
- Có kiến thức về Splunk, cơ chế bảo vệ web ModSecurity và cấu hình Apache.

1.1.4 Nội dung thực hành

- Trước khi khởi động bài lab, cần đảm bảo labtainer được cấu hình như sau:
    - Memory (RAM): 10GB
    - Hard Disk: Tối thiểu 80GB (khuyến nghị 100GB)
- Tải bài lab:

Vào /home/student/labtainer/labtainer-student và gõ lệnh sau trên terminal:

Imodule https://github.com/anhdnmit/do_an_tot_nghiep/raw/refs/heads/main/idr_splunk_webscan_new/imodule.tar

Khởi động bài lab:

labtainer -r idr_splunk_webscan

- Sau khi khởi động xong, 3 terminal ảo sẽ xuất hiện, một máy đại diện cho hệ thống splunk enterprise, một máy đại diện cho máy client thuộc splunk universal forwarder (UF), một máy đại diện cho máy attacker.

Để kiểm tra địa chỉ IP của 3 máy, gõ lệnh sau trên từng máy:

ifconfig

1.1.4.1 Triển khai và cấu hình hệ thống

1.1.4.1.1 Ở server – Triển khai và khởi tạo Splunk Enterprise

Đã tải sẵn bộ cài đặt Splunk ở client và server nên sinh viên chỉ cần giải nén và thực hiện lệnh. Trên server, sinh viên thực hiện lệnh sau:

sudo tar -xzf splunk-8.2.6-a6fe1ee8894b-Linux-x86_64.tgz -C /opt

sudo chown -R ubuntu:ubuntu /opt/splunk

Tiếp đến, lệnh ghi cấu hình OPTIMISTIC_ABOUT_FILE_LOCKING = 1 được thêm vào file splunk-launch.conf.

echo "OPTIMISTIC_ABOUT_FILE_LOCKING = 1" | sudo tee -a /opt/splunk/etc/splunk-launch.conf

sudo /opt/splunk/bin/splunk start --accept-license

Mục đích của dòng cấu hình này là yêu cầu Splunk bỏ qua kiểm tra khóa file trong môi trường ảo hóa hoặc container, vốn thường gây lỗi khi chạy trên nền tảng như Labtainer.

Cài đặt và đặt mật khẩu ngay trong quá trình khởi tạo xong thì thực hiện lệnh splunk status được dùng để bảo đảm dịch vụ splunkd đã hoạt động ổn định.

sudo /opt/splunk/bin/splunk status

Sinh viên sau đó mở trình duyệt (Firefox) tới địa chỉ http://127.0.0.1:8000, đăng nhập bằng tài khoản admin và mật khẩu đã đặt.

Ngay sau khi đăng nhập, một bước quan trọng khác là cấu hình Splunk lắng nghe dữ liệu gửi đến từ Universal Forwarder trên client.

sudo /opt/splunk/bin/splunk enable listen 9997 -auth <tài khoản>:<mật khẩu>

sudo netstat -tulpn | grep 9997

1.1.4.1.2 Ở client – Cài đặt Splunk Universal Forwarder (UF)

Phía client đóng vai trò nguồn gửi log về server nên cần triển khai Splunk Universal Forwarder.

sudo tar -xzf splunkforwarder-8.2.6-a6fe1ee8894b-Linux-x86_64.tgz -C /opt

sudo chown -R ubuntu:ubuntu /opt/splunkforwarder

Rồi đặt mật khẩu quản trị riêng :

sudo /opt/splunkforwarder/bin/splunk start --accept-license

Lệnh splunk status sau đó được dùng để kiểm tra xem UF đã khởi động đầy đủ hay chưa:

sudo /opt/splunkforwarder/bin/splunk status

Từ đó, ta thiết lập kết nối giữa UF và Splunk Enterprise:

sudo /opt/splunkforwarder/bin/splunk add forward-server 176.40.0.5:9997 -auth <tài khoản>:<mật khẩu>

1.1.4.1.3 Thiết lập log trên client – Tạo index trên server

Splunk sử dụng kiến trúc "index-based storage", nghĩa là mọi log phải được định danh vào một index cụ thể. Bởi vì bài lab tập trung phân tích tấn công vào web server, cần tạo một index riêng dành cho log web.

sudo /opt/splunk/bin/splunk add index web_lab -auth <tài khoản>:<mật khẩu>

tạo ra không gian lưu trữ độc lập web_lab, giúp việc truy vấn, lọc 404/403, thống kê IP tấn công… được tách biệt khỏi các log hệ thống khác. Khi index này tồn tại, Splunk Server đã sẵn sàng làm điểm nhận dữ liệu từ client.

Universal Forwarder (UF) trên client thực thi việc *thu thập log gốc* và *gửi đến server*. UF chỉ forward những file được khai báo trong inputs.conf. Vì vậy bước quan trọng nhất ở đây là hướng UF đến file log web Apache (access.log). Trong đó nội dung file của inputs.conf:

```
[monitor:///var/log/apache2/access.log]
sourcetype = access_combined
index = web_lab
disabled = false
```

Mọi thay đổi trong inputs.conf chỉ có hiệu lực sau khi restart UF, giúp UF nạp lại cấu hình, thiết lập session với server và bắt đầu forward dữ liệu.

Bước tiếp theo là chạy web để Apache sinh log mới.

Khi truy cập thử:

http://176.40.0.7/

http://176.40.0.7/login.php

➔ Apache tạo ra các dòng mới trong /var/log/apache2/access.log.

1.1.4.2 Phát hiện và phản ứng sự cố

1.1.4.2.1 Ở server : Phát hiện sự cố trên Splunk Server

Attacker sử dụng script webscan.sh, trong đó các URL được sinh ra dựa trên wordlist các đường dẫn phổ biến (admin, backup, phpmyadmin, test, …).

./webscan.sh http://176.40.0.7

Quá trình phát hiện diễn ra dựa trên việc phân tích log đã được forward vào index web_lab. Sinh viên truy cập Splunk http://176.40.0.5:8000 để kiểm tra log đã đổ vào index đó. Vào Search & Reporting, ô tìm kiếm được dùng để truy vấn index . Ta tra với lệnh cơ bản:

index=web_lab sourcetype=access_combined

Khi attacker đang thực hiện web scan, số lượng event tăng đột ngột trong realtime, thể hiện hiện tượng log tăng mạnh

Vì tấn công brute-force vào thư mục thường dẫn đến số lượng lớn request đến các đường dẫn không tồn tại, truy vấn lọc lỗi 404/403 giúp nhìn rõ hành vi:

index=web_lab sourcetype=access_combined status=404 OR status=403

Để chuyển từ dấu hiệu sang nhận diện chính xác nguồn tấn công, Splunk sử dụng SPL dạng thống kê:

```
index=web_lab sourcetype=access_combined
| stats count by clientip
| sort - count
```

Truy vấn này tổng hợp số lượng request theo từng IP, sắp xếp giảm dần. Nếu attacker là 176.40.0.3, địa chỉ này sẽ nổi bật với số lượng truy vấn vượt trội. Đây chính là dấu hiệu định lượng giúp xác định IP tấn công.

Bên cạnh đó, Splunk Server cũng cho phép truy vấn trực tiếp qua terminal. Lệnh:

sudo /opt/splunk/bin/splunk search index=web_lab sourcetype=access_combined -auth <tài khoản>:<mật khẩu> *-maxout 20 | tee -a evidence.txt*

Lưu các sự kiện đầu tiên vào file evidence.txt để phục vụ chứng cứ trong giai đoạn IR.

1.1.4.2.2 Ở client : Kiểm soát và chặn sự cố (Containment)

Khi xác định được IP tấn công, nhiệm vụ tiếp theo là ngăn chặn ngay lập tức để giảm thiểu ảnh hưởng. Bài lab sử dụng ModSecurity, một Web Application Firewall (WAF) dạng module của Apache. ModSecurity hoạt động theo cơ chế rule-based và có khả năng chặn request ngay trong pha xử lý đầu tiên của Apache.

Ta cần cài đặt và kích hoạt ModSecurity:

sudo apt-get update && sudo apt-get install -y libapache2-mod-security2

sudo a2enmod security2

sudo systemctl reload apache2

Khi module được kích hoạt, Apache có khả năng đọc rule từ file ModSecurity và áp dụng khi có request đến.

Mặc định ModSecurity chạy ở chế độ "DetectionOnly", chỉ ghi log mà không chặn. Vì vậy cần chuyển sang chế độ chủ động chặn bằng:

sudo nano /etc/modsecurity/modsecurity.conf

Rồi sửa sang: SecRuleEngine On

Đồng thời, thêm rule chặn IP tấn công

sudo nano /etc/modsecurity/custom_rules.conf

với nội dung:

SecRule REMOTE_ADDR "@ipMatch 176.40.0.3"
"id:1000001,phase:1,deny,log,status:403,msg:'Block 176.40.0.3'"

Giải thích:

- phase:1: chặn ở giai đoạn đầu tiên (before request body processing)
- deny,status:403: trả lại HTTP 403 Forbidden
- log: ghi log vào error.log
- ipMatch: so khớp theo địa chỉ IP chính xác

Để rule có hiệu lực, nó phải được include trong file cấu hình security2.conf

➔ thêm : IncludeOptional /etc/modsecurity/custom_rules.conf

Từ đó, khởi động lại Apache để ModSecurity được load và rule đã vào trong pipeline xử lý request.

1.1.4.2.3 Ở client : Diệt bỏ nguyên nhân (Eradication)

Khi quét log, người quản trị nhận ra đường dẫn /backup vốn để lộ thông tin nhạy cảm đã bị quét trúng.

Thay vì chặn một IP cố định, biện pháp eradication thiết kế lại cấu hình Apache để chỉ cho phép IP nội bộ truy cập đường dẫn nhạy cảm. File cần chỉnh:

sudo nano /etc/apache2/sites-available/000-default.conf

Thêm vào VirtualHost:

```
<Directory "/var/www/html/backup">
    Require ip 127.0.0.1 176.40.0.7
</Directory>
<Location "/backup">
    Require ip 127.0.0.1 176.40.0.7
</Location>
```

Chỉ máy chủ và chính client được quyền truy cập và bất kỳ IP nào khác (bao gồm attacker) đều bị chặn.

1.1.4.2.4 Trên client và server– Thu thập chứng cứ

Sau khi containment và eradication, cần thu thập bằng chứng phục vụ quá trình IR. Các bản log được đóng gói để đảm bảo tính nguyên vẹn:

mkdir -p ~/ir-backup

sudo tar czf ~/ir-backup/webscan_logs.tar.gz /var/log/apache2/access.log /var/log/apache2/error.log

sha256sum ~/ir-backup/webscan_logs.tar.gz > ~/ir-backup/webscan_logs.sha256

1.1.4.3 Theo dõi hậu sự cố

Trước tiên, sinh viên truy cập Search & Reporting để chạy một truy vấn chuyên biệt nhằm mô tả chính xác dấu hiệu hậu sự cố.

```
index=web_lab sourcetype=access_combined (status=404 OR status=403)
| stats count AS err_count by clientip
| where err_count >= 200
| sort - err_count
```

Chọn “Save As → Alert”, Splunk mở giao diện cấu hình Alert.

- Phần thông tin Alert
    - Title: WEB_SCAN_404_403_Alert
    - Description: mô tả mục tiêu <tùy chọn>
    - Quyền truy cập có thể để mặc định (Private).
    - Alert Type chọn Scheduled
    - Lịch chạy: Run on Cron Schedule, điền ô nhập : * * * * *

- Phần Trigger (điều kiện bắn cảnh báo)
    - Trigger alert when: Number of Results → is greater than: 0
    - Chế độ Trigger: Once – mỗi lần Splunk chạy truy vấn mà thấy kết quả, sẽ bắn một alert.
- Phần Actions (thao tác khi alert kích hoạt)
    - Event: WEB_SCAN detected: ip=$result.clientip$ count=$result.err_count$
    - Sourcetype: web_scan_alert
    - Index: web_lab

Để kiểm tra alert vận hành đúng, attacker tạo lại tấn công bằng script.

Trên giao diện web, truy vấn:

index=web_lab sourcetype=web_scan_alert

hoặc kiểm tra bằng CLI:

sudo /opt/splunk/bin/splunk search 'index=web_lab sourcetype=web_scan_alert' -auth <*tài* khoản>:<mật khẩu> | *tee -a evidence.txt*

1.1.4.4 Kết thúc bài lab:

Trên terminal đầu tiên sử dụng câu lệnh sau để kết thúc bài lab:

stoplab

Khi bài lab kết thúc, một tệp zip lưu kết quả được tạo và lưu vào một vị trí được hiển thị bên dưới stoplab. Trong quá trình làm bài sinh viên cần thực hiện lại bài lab, dùng câu lệnh:

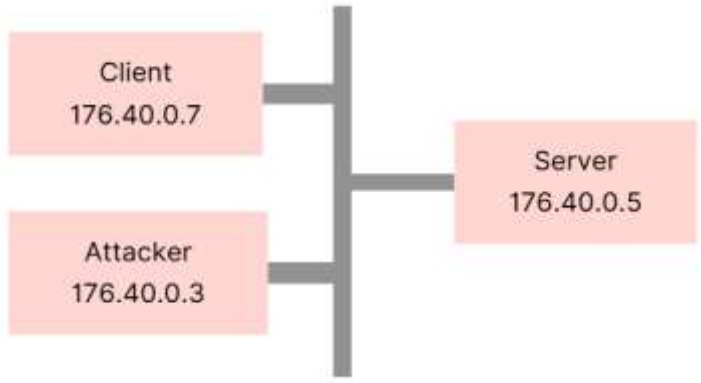
labtainer -r idr_splunk_webscan

1.1.5 Thiết kế bài thực hành

1.1.5.1 Cấu hình docker

- Trong môi trường máy ảo Ubuntu, sử dụng docker tạo ra 3 container: : 1 container “client”, 1 container “server, 1 container “attacker”.
- Tạo mạng có cấu hình:
    - Subnet: 176.40.0.0/24
    - External Gateway: 176.40.0.1

- Cấu hình docker gồm có:
    - Server: Cấu hình cho máy server
        - Tên máy: server
        - Địa chỉ trong mạng LAN: 176.40.0.5
        - Gateway: 176.40.0.1
    - Client: Cấu hình cho máy client
        - Tên máy: client
        - Địa chỉ trong mạng LAN: 176.40.0.7
        - Gateway: 176.40.0.1
    - Attacker: Cấu hình cho máy tấn công 2 sự cố
        - Tên máy: attacker
        - Địa chỉ trong mạng LAN: 176.40.0.3
        - Gateway: 176.40.0.1



Hình 1 Topo mạng bài idr_splunk_webscan

- config: lưu cấu hình hoạt động của hệ thống
- dockerfiles: mô tả cấu hình của các container, gồm:
    - server: sử dụng thư viện mặc định hệ thống, cập nhật hệ thống và cài sẵn Splunk Enterprise 8.2.6.
    - client: sử dụng thư viện mặc định hệ thống, cập nhật hệ thống và cài sẵn Splunk Universal Forwarder, Apache và môi trường web PHP.
    - attacker: sử dụng thư viện mặc định hệ thống, cập nhật hệ thống và cài sẵn script/webscan để thực hiện brute-force quét thư mục.

1.1.5.2 Các nhiệm vụ cần thực hiện

- Vận hành Splunk server để thu thập và phân tích log tập trung.
- Vận hành Splunk Forwarder trên client để gửi log web về server.
- Ghi nhận địa chỉ IP nguồn tấn công trong file bằng chứng trên server.
- Ghi nhận sự kiện truy cập trái phép bị web server chặn với mã lỗi 403.
- Cấu hình bảo vệ thư mục sao lưu `/backup` trên web server khỏi truy cập trái phép.
- Sao lưu dữ liệu liên quan thư mục backup và kiểm tra tính toàn vẹn bằng mã băm.
- Đối chiếu địa chỉ IP tấn công giữa log trên client và bằng chứng trên server để phân tích.

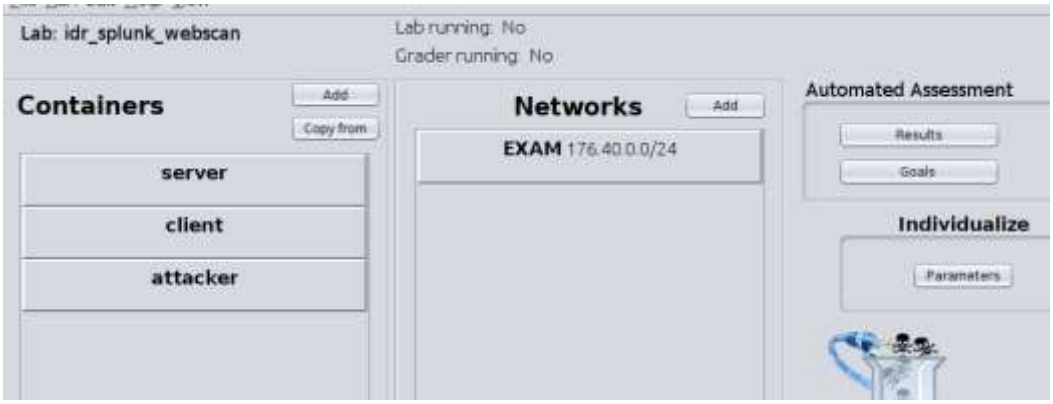
1.1.5.3 Các kết quả cần đạt được

Mỗi nhiệm vụ nhỏ sẽ được chia ra thành các mục chấm điểm để xác nhận sinh viên đã làm đúng các bước và hoàn thành bài thực hành hay chưa. Vì vậy, hệ thống sẽ ghi nhận các thao tác, sự kiện được mô tả theo bảng dưới đây để chấm điểm cho sinh viên:

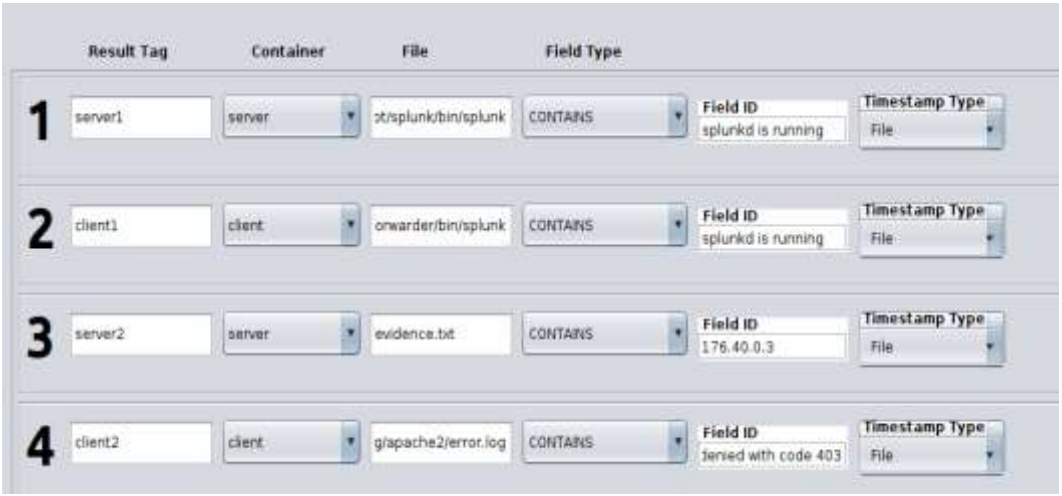
Bảng 1. Bảng result của bài idr_splunk_webscan

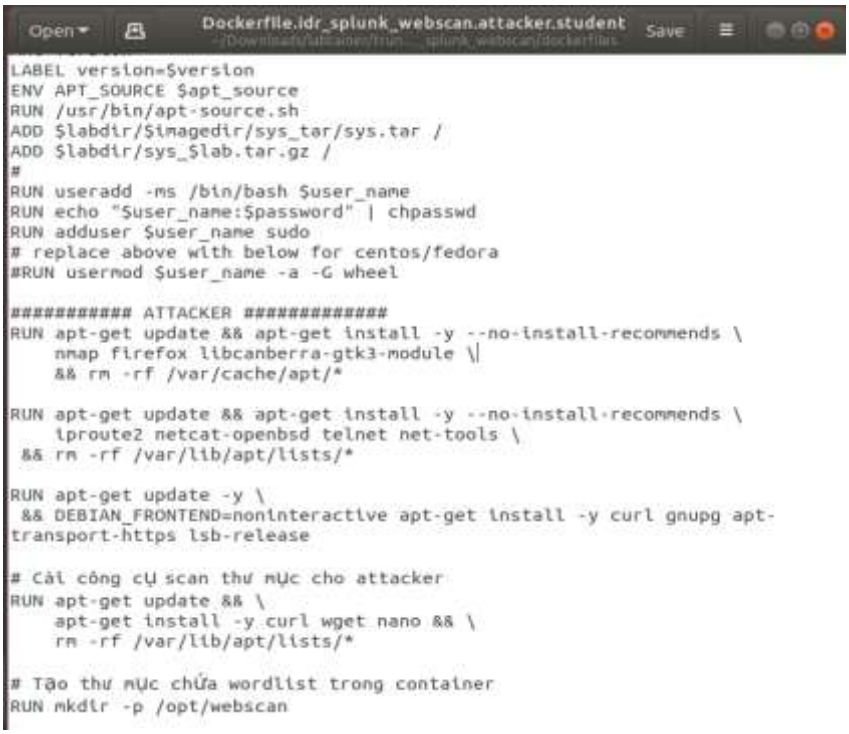
Result Tag	Container	File	File Type	Field ID	Timestamp Type
server1	server	/opt/splunk/bin/splunk	CONTAINS	splunkd is running	File
client1	client	/opt/splunkforwarder/bin/splunk	CONTAINS	splunkd is running	File
server2	server	evidence.txt	CONTAINS	176.40.0.3	File
client2	client	/var/log/apache2/error.log	CONTAINS	Access denied with code 403	File
client3	client	/etc/apache2/sites-available/000-default.conf	CONTAINS	<Directory "/var/www/html/backup">	File
client4	client	sha256sum.stdin	CONTAINS	backup	File
server3	server	evidence.txt	CONTAINS	ip=176.40.0.3	File

1.1.6 Cài đặt và cấu hình các máy ảo



Hình 2 Giao diện lab idr_splunk_webscan





Hình 3 Dockerfiles của máy attacker

Hình 4 Dockerfiles của máy client

Hình 5 Dockerfiles của máy server

1.1.7 Tích hợp và triển khai

1.1.7.1 Docker Hub

Hình 6 Đẩy lên docker hub

1.1.7.2 Github

Link github:

https://github.com/anhdnmit/do_an_tot_nghiep/tree/main/idr_splunk_portscanning

Hình 7 Đẩy file imodule.tar lên github

1.1.8 Thử nghiệm và đánh giá

Bài thực hành đã được xây dựng thành công. Bây giờ ta sẽ đánh giá về 1 trong các nhiệm vụ bài lab yêu cầu.

Trước hết, ta kiểm tra ip trên 3 máy.

```
ubuntu@client:~$ ifconfig
eth0: flags=4163<UP,BROADCAST,RUNNING,MULTICAST>  mtu 1500
        inet 176.40.0.7  netmask 255.255.255.0  broadcast 176.40.0.255
        ether 02:42:b0:28:00:07  txqueuelen 0  (Ethernet)
        RX packets 2695  bytes 216192 (216.1 KB)
        RX errors 0  dropped 0  overruns 0  frame 0
        TX packets 1854  bytes 743800 (743.8 KB)
        TX errors 0  dropped 0 overruns 0  carrier 0  collisions 0
```

Hình 8 IP client

```
ubuntu@server:~$ ifconfig
eth0      Link encap:Ethernet  HWaddr 02:42:b0:28:00:05
          inet addr:176.40.0.5  Bcast:176.40.0.255  Mask:255.255.255.0
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:357 errors:0 dropped:0 overruns:0 frame:0
          TX packets:280 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:0
          RX bytes:476391 (476.3 KB)  TX bytes:18167 (18.1 KB)
```

Hình 9 IP server

```
ubuntu@attacker:~$ ifconfig
eth0      Link encap:Ethernet  HWaddr 02:42:b0:28:00:03
          inet addr:176.40.0.3  Bcast:176.40.0.255  Mask:255.255.255.0
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:1632 errors:0 dropped:0 overruns:0 frame:0
          TX packets:2347 errors:0 dropped:0 overruns:0 carrier:0
          collisions:0 txqueuelen:0
          RX bytes:283829 (283.8 KB)  TX bytes:189695 (189.6 KB)
```

Hình 10 IP attacker

Ngoài 2 nhiệm vụ liên quan về mặt "Triển khai cấu hình" ở trên server và client, ở đây ta sẽ đánh giá nhiệm vụ "Ghi nhận địa chỉ IP nguồn tấn công trong file bằng

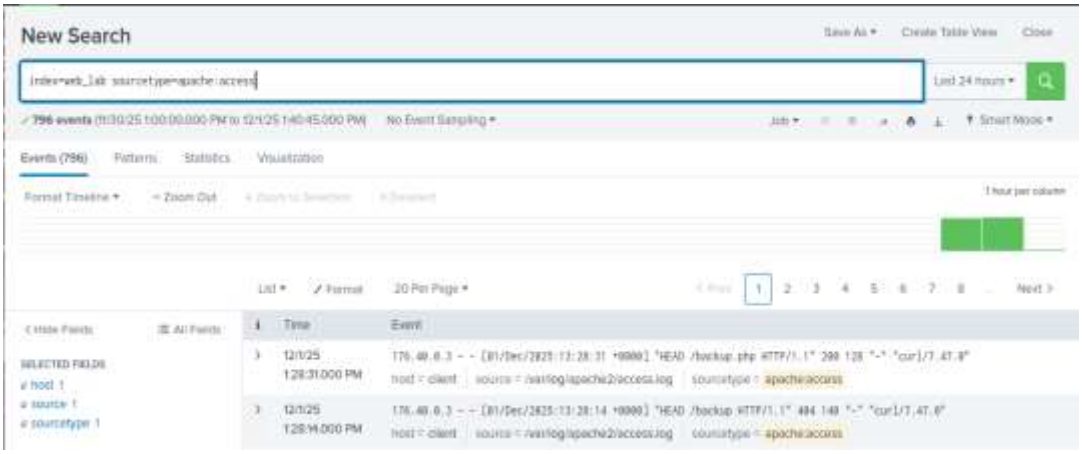
chứng trên server" trong giai đoạn "Phát hiện sự cố", liên quan đến các nhiệm vụ tiếp theo của hướng dẫn.

Toàn bộ bước làm xoay quanh việc dùng Splunk để phát hiện IP tấn công, rồi dùng lệnh Splunk CLI trên server để lưu kết quả vào file.

Trước hết, ta mở giao diện Splunk trên server để quan sát log web. Trên trình duyệt, truy cập địa chỉ: http://176.40.0.5:8000

Sau khi đăng nhập, để bắt đầu phân tích log, ta chuyển sang ứng dụng tìm kiếm. Trên thanh Apps, chọn: Apps → Search & Reporting

Lúc này giao diện Search sẽ xuất hiện, cho phép ta nhập câu lệnh truy vấn log. Tiếp theo, ta cần kiểm tra xem log web đã được đẩy vào đúng index hay chưa. Tại ô Search, ta nhập câu truy vấn cơ bản:



Hình 11 Giao diện tìm kiếm

Kỳ vọng ở đây là Splunk sẽ trả về các event từ file access.log, với các trường như: clientip, status, uri_path, _time, method. Khi attacker thực hiện quét web, số lượng event trong index này sẽ tăng rất nhanh, ta có thể thấy rõ trên màn hình Search. Từ kết quả đó, ta bắt đầu thu hẹp phạm vi bằng làm chặt lại điều kiện kiểm tra để tập trung vào dấu hiệu bất thường.

Hình 12 Checkwork bài idr_splunk_webscan