Log File Analysis Report

Student Name: Jerome Arsany Mansour Farah

Id: 2305093 (Level 2)

Course: Information Security Mangement

Introduction

This project analyzes a web server access log file using a Bash script to extract key metrics, detect failures, and identify usage patterns. It was developed on Kali Linux as part of an information security management course. The goal is to gain insights into server access behavior, identify potential issues, and propose improvements.

Objective

Analyze a web server access log file using Bash to identify request patterns, errors, and potential anomalies.

Tools Used

- Kali Linux
- Bash scripting
- awk, grep, sort, uniq
- Sample Apache log file

Files Submitted

- access.log (Apache log file)
- log_analysis.sh (Bash script)
- report.pdf (this document)

Bash Script Code

```
#!/bin/bash

LOG_FILE="access.log"

echo "Total Requests:"
wc -l < "$LOG_FILE"

echo "GET Requests:"
grep -c 'GET' "$LOG_FILE"

echo "POST Requests:"
grep -c 'POST' "$LOG_FILE"

echo "Unique IPs:"
awk '{print $1}' "$LOG_FILE" | sort | uniq | wc -l

echo "GET/POST by IP:"
awk '$6 ~ /GET|POST/ {print $1, $6}' "$LOG_FILE" | sort | uniq -c | sort -nr

echo "Failed Requests:"
FAILS=$(awk '$9 ~ /^[45]/' "$LOG_FILE" | wc -l)
TOTAL=$(wc -l < "$LOG_FILE")
PERCENT=$((FAILS * 100 / TOTAL))
echo "Failures: $FAILS"
echo "Failure Rate: $PERCENT%"

echo "Most Active IP:"
awk '{print $1}' "$LOG_FILE" | sort | uniq -c | sort -nr | head -1

echo "Daily Request Average:"
awk -F'[:[]' '{print $2}' "$LOG_FILE" | cut -d/ -f1,2,3 | sort | uniq -c | awk '{total+=$1;
count++} END {print total/count}'

echo "Failure Requests by Day:"
awk '$9 ~ /^[45]/ {split($4, a, ":"); split(a[1], d, "/"); print d[1] "/" d[2] "/" d[3]}'
"$LOG_FILE" | sort | uniq -c | sort -nr | head

echo "Requests Per Hour:"
```

```
awk -F'[:[]' '{print $2":"$3}' "$LOG_FILE" | cut -d: -f1-2 | sort | uniq -c

echo "Status Code Breakdown:"
awk '{print $9}' "$LOG_FILE" | sort | uniq -c | sort -nr

echo "Most Active IP for GET:"
grep 'GET' "$LOG_FILE" | awk '{print $1}' | sort | uniq -c | sort -nr | head -1

echo "Most Active IP for POST:"
grep 'POST' "$LOG_FILE" | awk '{print $1}' | sort | uniq -c | sort -nr | head -1

echo "Failure Patterns by Hour:"
awk '$9 ~ /^[45]/ {split($4, a, ":"); print a[2]}' "$LOG_FILE" | sort | uniq -c | sort -nr
```

Sample Log File (access.log)

The access log file (`access.log`) is included in the GitHub repository for review. Due to its size and format, it is not printed here but can be opened with any text editor or analyzed using the Bash script.

Conclusion

Through this project, various metrics such as total requests, failed requests, and IP-based activity were extracted from a server log. This analysis helps in identifying usage trends and potential vulnerabilities. Future improvements could include automating alert systems for failure spikes or unusual IP behavior.