Log File Analysis Report

Alexandria National University

Faculty of Computer And Data Science

Department of Cyber Security

💻 Log File Analysis Report

Prepared by: Hanin Mohamed Hamouda
ID: 2205232
Course: Information Security
Instructor: Yahya Ashraf
Submission Date: 9 May 2025

1. Executive Summary

This document provides an in-depth analysis of server access logs using a Bash script. The analysis includes request volumes, user behavior patterns, and error breakdowns, and concludes with actionable recommendations to improve system performance and reliability.

2. Objective

The objective is to automate the analysis of server logs to extract valuable metrics, detect anomalies, and improve the overall understanding of web traffic and system behavior. This includes identifying request trends, high-traffic users, failure rates, and potential system weaknesses.

3. Analysis Overview

Top 5 Most Active IPs

Rank	IP Address	Requests
1	66.249.73.135	482
2	46.105.14.53	364
3	130.237.218.86	357
4	75.97.9.59	273
5	50.16.19.13	113

Average Daily Requests: 2500.00

Failures per Day

Date	Failures
17/May/2015	30
18/May/2015	66
19/May/2015	66
20/May/2015	58

Requests per Hour

Hour	Requests
10	443
11	459
12	462
13	475
14	498
15	496
16	473
17	484
18	478
19	493
20	486
21	453

22	346
23	356

Requests per Day (Trend)

Date	Requests
17/May/2015	1632
18/May/2015	2893
19/May/2015	2896
20/May/2015	2579

Status Code Breakdown

Status Code	Description	Count
200	OK	9126
206	Partial Content	45
301	Moved Permanently	164
304	Not Modified	445
403	Forbidden	2
404	Not Found	213
416	Range Not Satisfiable	2
500	Internal Server Error	3

Top Users by Request Method

- Top GET IP: 66.249.73.135 — 482 GET requests

- Top POST IP: 78.173.140.106 — 3 POST requests

Failure Requests per Hour

Hour	Failures
10	12
11	11
12	7
13	12
14	11
15	6
16	8
17	12
18	9
19	10
20	4
21	8
22	8
23	4

4. Key Findings

- The IP 66.249.73.135 had the highest number of requests and appears to be a crawler.

- Server traffic was concentrated between 10:00 and 20:00 with multiple peak hours.

- Errors are most common during peak hours, indicating load issues or misconfigurations.

- The majority of failed requests were 404 errors, likely from broken links or bot scanning.

- Very few POST requests were made, suggesting limited interactive traffic.

5. Recommendations

- Investigate the top 5 IP addresses (e.g., 66.249.73.135, 46.105.14.53) for potential scraping, bot activity, or abusive acces

- Review peak request hours (especially between 10:00–20:00) to ensure server resources are sufficient and response times are stable.

- Apply rate limiting or implement CAPTCHA challenges for suspicious or high-frequency IPs.

- Review and update broken or outdated links causing 404 errors. This may also improve SEO and user experience.

- Investigate all 5xx status codes (even if few) to identify any server-side failures or misconfigurations.

- Consider load balancing if traffic continues to increase during peak periods.

- Analyze low POST request count — this may indicate missing form submissions or read-only behavior where write operations are expected.

- Schedule regular log analysis (weekly or daily) to detect anomalies or usage pattern changes over time.

- Implement centralized logging and monitoring tools (e.g., ELK Stack, Splunk) for long-term insights and alerting.

- Review firewall rules and access controls for repeated failed access attempts from specific IPs.

6. Conclusion

This log analysis provides a comprehensive view of server usage and request behavior. By applying the recommended improvements, the system’s reliability, performance, and security can be enhanced significantly.