Log File Analysis Report

Analysis by Ahmed Hassan

Saturday 10th May, 2025

Contents

1 Introduction

This report provides a comprehensive analysis of the access.log file, generated on Saturday, May 10, 2025, at 09:44:05 EET (Egypt local time). The analysis includes request counts, unique IP addresses, failure rates, daily request averages, and patterns in request behavior. Suggestions for system improvements and security measures are provided based on the findings.

2 Request Counts

- **Total Requests**: 9344
- **GET Requests**: 9311
- **POST Requests**: 0

3 Unique IP Addresses

- **Total Unique IPs**: 636
- **GET and POST Requests per IP**: Detailed counts for each IP are provided in the appendix (Table **??**).

4 Failure Requests

- **Total Failed Requests (4xx, 5xx)**: 1533
- **Percentage of Failed Requests**: $\frac{1533}{9344} \times 100 \approx 16.40\%$

5 Top User

- **Most Active IP**: 82.154.31.44 (338 requests)

6 Daily Request Averages

- **Average Requests per Day**: Data spans from June 2, 2010, to May 24, 2011 (356 days). Average requests per day: $\frac{9344}{356} \approx 26.25$

7 Failure Analysis

- **Days with Highest Failure Requests**:
  1. 26/May/2011: 109 failures
  2. 03/May/2011: 81 failures
  3. 01/Feb/2011: 33 failures
  4. 02/Feb/2011: 24 failures
  5. 28/Mar/2011: 23 failures

8 Requests by Hour

- **Requests per Hour**: See Table 1.

Table 1: Requests by Hour

Hour	Requests
00	238
01	257
02	85
03	62
04	81
05	136
06	76
07	93
08	174
09	473
10	280
11	609
12	983
13	407
14	736
15	574
16	518
17	658
18	584
19	319
20	726
21	521
22	368
23	387

9 Status Codes Breakdown

- **Status Code Frequencies**:
  - Status 200: 6294 occurrences
  - Status 304: 1505 occurrences
  - Status 404: 851 occurrences
  - Status 403: 678 occurrences
  - Status 206: 11 occurrences
  - Status 401: 4 occurrences
  - Status 301: 2 occurrences

10 Most Active User by Method

- **Most Active GET IP**: 82.154.31.44 (338 GET requests)
- **Most Active POST IP**: None (0 POST requests)

11 Patterns in Failure Requests

- **Failure Requests by Hour**: See Table 2.

Table 2: Failure Requests by Hour

Hour	Failures
00	45
01	33
02	25
03	37
04	34
05	46
06	29
07	30
08	45
09	76
10	37
11	61
12	182
13	52
14	92
15	71
16	77
17	87
18	157
19	63
20	70
21	55
22	52
23	77

12 Request Trends

- **Peak Request Hour**: Hour 12 with 983 requests
- **Lowest Request Hour**: Hour 03 with 62 requests
- **Trend**: Requests peak at hour 12, indicating high server load during midday, likely due to user activity or automated processes.

13 Analysis Suggestions

Based on the analysis, the following suggestions are proposed:

- **Reduce Failures**: The 1533 failed requests (16.40%) are primarily 403 (678) and 404 (851) errors. Review access controls to ensure legitimate requests are not blocked (403). Investigate missing resources or broken links causing 404 errors.
- **Monitor Peak Times**: Hour 12 (983 requests) and hour 18 (157 failures) show high activity and failures. Scale server resources or implement caching during these periods to handle load.
- **Security Concerns**: IP 82.154.31.44 made 338 requests, indicating potential bot or scraping activity. Other IPs (e.g., 188.140.124.209) show rapid requests. Implement rate-limiting and monitor for abuse.
- **System Improvements**: Deploy a Web Application Firewall (WAF) to filter malicious traffic. Optimize server for bot traffic (e.g., Googlebot, SheenBot) by ensuring proper robots.txt and crawl budgets.