NASA HTTP

Log File Analysis Report

Ahmed Sabry Ibrahim

2205007

Information Security

Management

Table of Contents

Executive Summary

The server processed 1,569,898 requests over 31 days with a low failure rate of 0.65%. Key observations include consistent traffic patterns, minimal POST requests, and critical security insights from active IPs.

Key Metrics

Total Requests	1,569,898
GET Requests	1,565,812 (99.74%)
POST Requests	111 (0.01%)
Unique IP Addresses	75,060
Failed Requests (4xx/5xx)	10,251 (0.65%)
Daily Average Requests	50,641.87 requests/day

Traffic Patterns

- **Hourly Distribution**
  - Peak Hours: 12:00–15:00 (~105,000–109,000 requests/hour)
  - Lowest Activity: 02:00–04:00 (~26,000–32,000 requests/hour)
- **Top High-Error Days**
  - 30/Aug/1995: 80,641 errors
  - 31/Aug/1995: 90,125 errors
  - 29/Aug/1995: 67,988 errors

Error Analysis

- **Most Common Errors**

Code	Count	Description
404	9,978	Not Found
403	171	Forbidden
500	6	Internal Server Error

- **Failure Trends**
    - **Critical Failure Periods:** 12:00–15:00 (500–600 errors/hour)
    - **Unexpected Spike:** 02:00 with 618 errors.

Security Observations

- **Top Active IPs**
    - **edams.ksc.nasa.gov** (6,530 requests)
    - **piweba4y.prodigy.com** (4,846 requests)
    - **163.206.89.4** (4,791 requests)
- **Anomalies**
    - **Suspicious Activity:** Repeated 404 errors for **/admin.php** (239 failures).
    - **High-Risk IPs: www-d1.proxy.aol.com** (3,889 requests) – monitor for potential scraping.

Recommendations

1. **Error Reduction**
   - Fix broken links causing 404 errors (e.g., /admin.php).
   - Investigate 500 errors via server logs.
2. **Performance Optimization**
   - Scale server capacity during peak hours (12:00–15:00).
   - Cache static assets (e.g., images, CSS/JS files).
3. **Security Enhancements**
   - Block IPs with >5,000 requests/day.
   - Implement rate limiting for /admin* endpoints.
4. **Monitoring**
   - Set alerts for 500 errors and traffic spikes.
   - Audit IPs with high request volumes for malicious intent.

Conclusion

The server demonstrates strong performance under high traffic, but proactive measures are needed to address security risks and optimize error handling. Further analysis of peak-hour failures and suspicious IPs is recommended.

Attachments & Referenes

- Full Script: analyze.sh
- Analysis Report: server_analysis.txt
- Log Data: access.log (Download)
    - Log Data Reference: M. Arlitt and C. Williamson, entitled ``Web Server Workload Characterization: The Search for Invariants", to appear in the proceedings of the 1996 ACM SIGMETRICS Conference on the Measurement and Modeling of Computer Systems, Philadelphia, PA, May 23-26, 1996