

# Apache Log Analysis Report

## 1. Introduction

This report analyzes an Apache web server log file using a Bash script to extract key metrics on requests, IPs, and failures.

The goal is to identify patterns, detect issues, and suggest improvements for server performance and security.

The log covers 10,000 requests over 4 days.

## 2. Statistics

### 1. Request Counts:

- Total: 10,000
- GET: 9,952 (99.52%)
- POST: 5 (0.05%)

### 2. Unique IP Addresses:

- Most Active: 66.249.73.135 (482 GET requests, 4.82%).

### 3. Failed Requests:

- 4xx/5xx: 220
- Failure Rate: 2.00%

### 4. Most Active User:

- IP: 66.249.73.135 (482 requests).

### 5. Daily Request Average:

- 2,500 requests/day (4 days).

#### 6. Failure Analysis by Day:

- May 18: 66 failures
- May 19: 66 failures
- May 20: 58 failures

#### 7. Requests by Hour:

- Highest: 14:00 (498 requests)
- Lowest: 08:00 (345 requests)
- Peak: 12:00-20:00

#### 8. Request Trends:

- Peak Hour: 14:00 (498 requests)
- Trend: Afternoon activity

#### 9. Status Code Breakdown:

- 200 (OK): 9,126
- 304 (Not Modified): 445
- 404 (Not Found): 213
- 301 (Redirect): 164
- 206 (Partial Content): 45
- 500 (Server Error): 3
- 416 (Range Not Satisfiable): 2
- 403 (Forbidden): 2

#### 10. Most Active IP by Method:

- GET: 66.249.73.135 (482)
- POST: 78.173.140.106 (3)

#### 11. Failure Patterns:

- Top Hours: 09:00 (18 failures), 05:00 (15), 06:00 (14)

#### 3. Analysis

- GET requests dominate (99.52%), indicating static content.
- 66.249.73.135 (likely Googlebot) is the most active IP.
- 2% failure rate, including 213 404 errors and 3 500 errors.
- High failures on May 18-19 may be due to maintenance or load.
- Peak traffic occurs in the afternoon, especially at 14:00.
- Most failures happen early morning (05:00-09:00).

#### 4. Suggestions

##### 1. Reduce Failures:

- Fix 404 errors using `wget --spider`.
- Check error log: `/var/log/apache2/error.log`

##### 2. Monitor High-Failure Days:

- Review logs for May 18-19.

##### 3. Handle Peak Hours:

- Scale resources for 12:00-20:00, especially 14:00.

##### 4. Security:

- Rate-limit IP 66.249.73.135 (robots.txt)

- Secure forms using CSRF and CAPTCHA

#### 5. Improvements:

- Use caching (Varnish/Cloudflare)
- Deploy WAF for early morning traffic
- Monitor with Prometheus/Nagios

#### 5. Conclusion

The Apache log analysis reveals a stable server with a 2% failure rate. GET requests dominate the traffic.

Fixing 404 and 500 errors, optimizing peak-hour performance, and improving monitoring and security will enhance reliability.