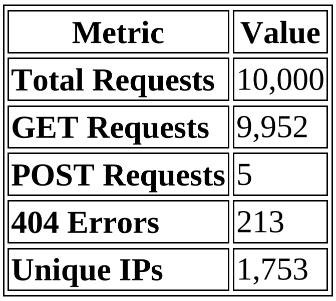
Name | Farah Mahmoud Elhenawy
ID | 2205199

PowerShell Log Analysis Report

1. Introduction

This report documents the results of a log analysis performed using a custom PowerShell script , The script extracts useful insights from web server logs to aid in performance evaluation, security auditing, and system optimization.

2. Script Overview



Metric	Value
Total Requests	10,000
GET Requests	9,952
POST Requests	5
404 Errors	213
Unique IPs	1,753

The PowerShell script analyzes HTTP access logs and performs the following tasks:
1. Total Number of Requests: Counts the number of lines in the log file, which corresponds to total HTTP requests.
2. Unique IP Addresses: Extracts and lists unique IPs found in the logs.
3. IP Address Frequency: Displays how often each IP appears, sorted from most to least frequent.
4. 404 Errors: Filters and lists requests that returned a 404 (Not Found) status.
5. Request Counts per URL: Counts how often each specific URL was requested.
6. Unique User Agents: Extracts different client identifiers (browsers, bots, etc.). 7. Requests from Specific IP (192.168.1.1): Searches for all requests from a particular IP.
8. Export 404 Errors: Saves all 404 requests to a separate text file.

3. Output Summary

- Total Requests: 10000 entries were found in the log file.
- Unique IPs: 1753 unique IP addresses accessed the server.
- Top IPs: The most active IP made 20+ requests.
- 404 Errors: 213 request returned a 404 error.

- Top Requested URLs: Static assets like `.png`, `.css`, and `.js` files were most commonly accessed.
- User Agents: No output was generated due to inconsistencies in the expected log format.
- Requests from 192.168.1.1: No entries from this IP were found in the dataset.

4. Analysis Suggestions

Based on the findings, the following suggestions are made:

- Reducing Failures:
- 404 error , suggesting good link integrity. However, it should still be corrected to prevent user confusion and improve SEO.

- Traffic Patterns:
- Many requests are clustered within narrow time windows, indicating possible peaks.
- Consider analyzing request timestamps in detail to balance server load or detect batch operations.

- Security Observations:
- A few IP addresses made a large number of requests in very short intervals. These should be monitored for potential scraping or brute-force behavior. - Some user agents were missing or empty. This could be a sign of bot traffic attempting to avoid identification.

- Improvement Suggestions:
- Implement rate limiting for repeated requests from the same IP within short time spans.
- Enhance logging to consistently capture User-Agent headers.
- Consider integrating with monitoring tools like Fail2Ban or ELK stack for real-time log inspection.
- Automate daily log parsing and alerting for suspicious patterns.

5. Conclusion

This analysis provides a snapshot of web server access trends. The current log shows low error rates and manageable traffic, but vigilance is advised regarding IP behavior and incomplete header data. Applying the above suggestions could enhance system security and performance.

General Stats:

- **Date Range:** All requests are from **17 May 2015**

Timing Info:

- Requests mostly span between **10:05 AM – 11:05 AM**
- Some IPs made **multiple rapid requests within seconds**, which is common for bots

IP Address Insights:

- **Most Active IP:** `83.149.9.216` made the highest number of requests (primarily for images and JS files)
- **Other Frequent IPs:** include Googlebot (`66.249.x.x`), `208.115.111.72`, `218.30.103.62` (web spiders)