Network Traffic Analysis Report – I used the same logfile

Overview Statistics

- **Total Analyzed Requests**: 10,000
- **Request Method Distribution**: 9,952 GET requests (99.52%), 5 POST requests (0.05%)
- **Failure Rate**: 2.20%
- **Average Daily Traffic**: 2,500 requests/day

Request Analysis

Method Distribution Analysis

The overwhelming majority of traffic consists of GET requests (99.52%), with minimal POST activity (0.05%). This pattern indicates the website primarily serves content rather than processing user submissions or data entry. The lack of interactive submissions suggests a content-focused platform.

IP-based Request Patterns

During the analysis period (March 17-20), several IP addresses generated abnormally high GET request volumes with no corresponding POST activity. Most notably:

- IP 66.249.73.135 generated 482 GET requests
- Multiple other addresses exceeded 100 requests each

This concentrated activity pattern may indicate:

- Automated crawling behavior
- Potential scanning activity
- Possible unauthorized access attempts

Status Code Distribution

- **Successful responses**: >90% (status code 200)
- **Redirects/cached content**: Small number of 301 and 304 responses
- **Client errors**: 213 instances of 404 errors (missing resources)
- **Server issues**: Minimal 500-class errors indicating server failures

- **Access restrictions**: Several 403 responses suggesting permission issues

Temporal Analysis

Daily Request Distribution

1. May 19, 2015: 2,896 requests
2. May 18, 2015: 2,893 requests
3. May 20, 2015: 2,579 requests
4. May 17, 2015: 1,632 requests

Hourly Traffic Patterns

Traffic demonstrates clear diurnal patterns:

- **Peak activity**: 10:00-21:00, with maximum volume between 14:00-15:00
- **Minimal activity**: 00:00-09:00
- **Sharp decline**: After 22:00

Sample hourly counts:

107 requests at 18:00 on May 20
110 requests at 08:00 on May 18
110 requests at 17:00 on May 19
111 requests at 11:00 on May 17
111 requests at 23:00 on May 17
111 requests at 07:00 on May 19
112 requests at 15:00 on May 19
112 requests at 11:00 on May 20
112 requests at 12:00 on May 20
113 requests at 19:00 on May 18

Error Distribution

- **Peak error days**: May 18-19 (66 failures each)
- **Reduced errors**: May 17 (possibly due to lower overall traffic)
- **Notable error spike**: 09:00 on May 20

Security Assessment

Anomalous Activity

Several IP addresses exhibited request volumes significantly above the statistical norm:

- Most addresses generated fewer than 50 requests
- Select addresses exceeded 100 requests
- One address approached 500 requests (66.249.73.135, likely Googlebot)

These activity spikes occurred at various times and from different sources, suggesting potential:

- Crawler activity
- Vulnerability scanning attempts
- Resource exhaustion tactics

Recommendations

Error Reduction Strategies

1. Investigate peak failure periods (May 18-19) for:
    a. Recent code deployment issues
    b. System resource constraints
    c. Potential security incidents
2. Address 404 errors through:
    a. Comprehensive link validation
    b. Implementation of proper redirects for relocated resources
    c. Regular content auditing
3. Mitigate server errors by:
    a. Enhancing error handling procedures
    b. Implementing input validation
    c. Developing resilience during peak load periods

Infrastructure Optimization

1. Deploy rate limiting mechanisms to prevent resource exhaustion
2. Implement load balancing during peak hours (10:00-21:00)
3. Consider scaling adjustments based on temporal patterns
4. Optimize caching strategies to reduce backend processing requirements

Security Enhancements

1. Expand logging capabilities to include:
    a. Detailed IP information
    b. Endpoint access patterns
    c. User-agent analysis
2. Implement real-time alerting for:
    a. Error rate spikes
    b. Unusual request patterns
    c. Authentication anomalies
3. Audit authentication mechanisms to detect potential brute-force attempts
4. Review resource access controls to address 403 responses