

Engineering Sector Skills Project Report

By Hannan Nur

In this sector skills project, I have used Java and object oriented programming to identify the traffic issues in the podcast and newsletter platform and to develop a solution to handle the increased downtime in the server.

Assumptions

What I assumed in the client's web server access log while tackling the coding part of the project were:

- The access log format is standardised and consistent across each entry.
- The format includes: user's IP address, country code, time stamp, user request, HTTP status code, number of bytes used for the log, referrer, user agent and response time.

Findings

- I created a Java system with four classes:
 1. **WebLog**: a record that stores each log as an object with the IP address, country code, time stamp and user agent as its parameter.
 2. **LogLoader**: a parser that loads log entries into WebLog objects.
 3. **BotDetector**: a class that detects bot activities based on analysing the WebLog objects
 4. **Logs**: a class that prints out the details of the suspected bots
- I found some IP addresses made over 100 requests within a few minutes - e.g. IP address 194.168.1.1. This is an indicator of bot activity.
- The IP addresses came from various countries (NO, DE, RU, KP, DK, IR, CN, FR, SE, AU, UK, US, CA, NL) and often paired with identical generic user agents.

Cost Effective Solutions

- Block the IP addresses that exceeded a high volume of requests within a short time period.
- Use bot filters - e.g. CAPTCHA and Cloudflare.
- Implementing these solutions should reduce the traffic and the downtime in the server. Furthermore, this will increase the productivity of the small engineering team.