Bot Traffic Control Report

What’s Going On

Following a surge in traffic to the company’s podcast and newsletter platform, server performance issues have emerged mainly the frequent downtime. I write this report to outline the findings from a log analysis to identify potential bot traffic and offers cost effective recommendations suitable for a small startup with limited funding and engineering resources.

What I did

A Python script was developed to analyse server access logs containing 432091 entries. The script extracted main data points such as IP address, requested endpoint, and HTTP status code. In the script, three main indicators were used to flag suspicious IPs:

* High Request Volume: IPs that make more than 50 requests.
* Frequent 404 Errors: IPs that generate over 10 “404” errors.
* API Endpoint Access: IPs that access API / endpoints more than 5 times.

These thresholds were chosen to balance sensitivity and specificity, aiming to detect aggressive or automated traffic without flagging regular users.

What I found

The analysis showed a significant number of IPs exhibiting bot like behaviours. Many IPs accessed API endpoints repeatedly, generated excessive requests and errors, suggesting automated scraping and probing activity. These IPs are likely the reason of server strain and reduced availability.

My assumptions

* The current system doesn’t have bot protection in place.
* The log format is consistent and includes IPs, endpoints, and status codes.
* The company has a small team and low budget, so the solution needs to be simple and affordable.

Recommendations

* Adding basic rate limiting to slow down aggressive IPs.
* Blocking or monitoring IPs that show suspicious patterns.
* Protecting API endpoints with authentication or Captcha.
* Setting up a new lightweight monitoring script to flag future issues automatically.

Final Thoughts

This approach is low-cost, easy to maintain, and gives the team a clearer picture of what’s happening behind the scenes. It’s a solid first step toward keeping the site stable and bot-free.