#ASLI ENGINEERING

Rate Limiter took down GitHub

SWIPE

BY

ARPIT BHAYANI

Dissecting GitHub Outage

Downtime due to Rate limiter

Rate limiters are meant to avoid downtime,
but GitHub went down due to a rate limiter

What happened?

A large fraction of users saw errors while using GitHub

↓

The affected users were part of
an A/B experiment

What is A/B experiment? Data never lies

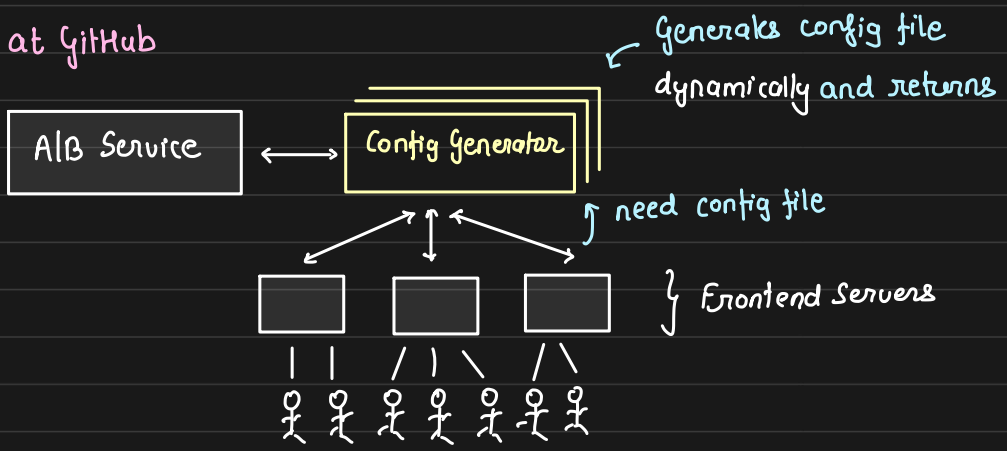
To quantify which variation is better,
we run A/B test.

A fraction of users are shown variation
while others are shown existing behaviour.

By observing key metrics by variation, we can
quantify if variation is positive or not.

When in doubt run an A/B experiment.

A|B at GitHub



A configuration file is dynamically generated and sent upon request to frontend servers. The config is used to run very granular A|B tests.

What failed ?

Because of a lot of request, the file was not generated for all the servers. Request was rate limited by the config generator. Hence all the users part of the variation group got affected !!

* Rate limiting (even internal) is essential, but limits should be tuned well !

Mitigation and long term fix

GitHub team quickly disabled their dependence on this file and this restored the services to all the users

long-term fix: Configuration file to be computed and cached and sent to all who need it instead of making runtime requests.

Key Takeways

- Avoid synchronous dependencies
  - ↳ make communication async wherever possible
- Rate limiters are great
  - ↳ but tune them well. Have higher limits for tier 1 services
- Create tier of services and classify your services depending on the criticality.