Experiments for SRE

Debbie Ma

Google

SRECon'21

What are experiments?

Ideation, Trial, Launch

Ideation

Will my new feature increase app usage?

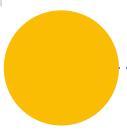
Trial

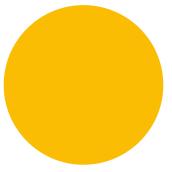
New feature rolled out to 10% of the population.

Launch

New feature saw increased app usage :)







Ideation, Trial, Launch

Ideation

Is the new feature **reliable**?

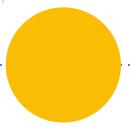
Trial

New feature rolled out to 10% of the population.

Launch

New feature saw increased crashes:(







How can SRE use experiments?

SRE best practices

Gradual Rollouts

Ramp experimental features

Change Attribution

Find changes associated with **experiments**.

Controlled Mitigation

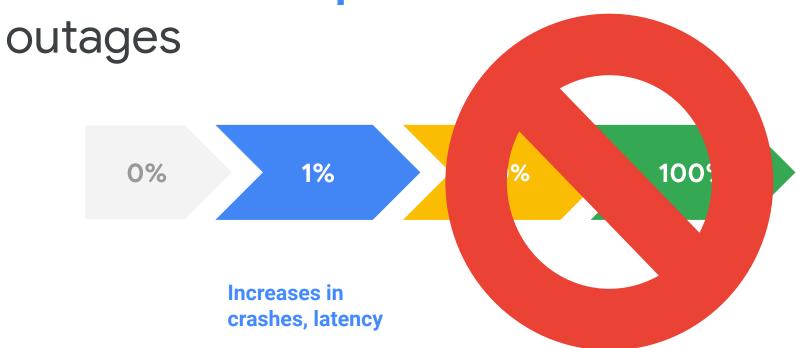
Rollback **experiments** guarding new code paths

Reduce the impact of outages



Increases in crashes, latency

Reduce the impact of



Root cause outages

- Experiment guarded features launched to production
- Experiment ID associated with unhealthy metrics

Mitigate outages



Increases in crashes, latency

"Rolling back the feature flag is as simple as ramping the launch back down to 0%, instead of rebuilding an entire binary with the fix to release to the world."

- Engineering Reliable Mobile Applications

Enforcing reliability with experiments

Reliable production changes

- Ramp changes in stages
- Ramp up only if the experiment is healthy
- Transition experiment change to permanent change

Engineer reliability into experiments

- Automate checks at every ramp stage
- Encode best practices into "policies"

Experiments for SRE tomorrow

Experiments as change management

- Enhance monitoring to improve outage attribution
- Develop more policies to automate away decisions
- Experiments as part of the canonical production change management system

Experiment infrastructure changes for mitigation

- Experiment signal speed?
- Reduce experiment learning curve

Thank You