goto;

GOTO
AARHUS 2021

#GOTOaar

About Me

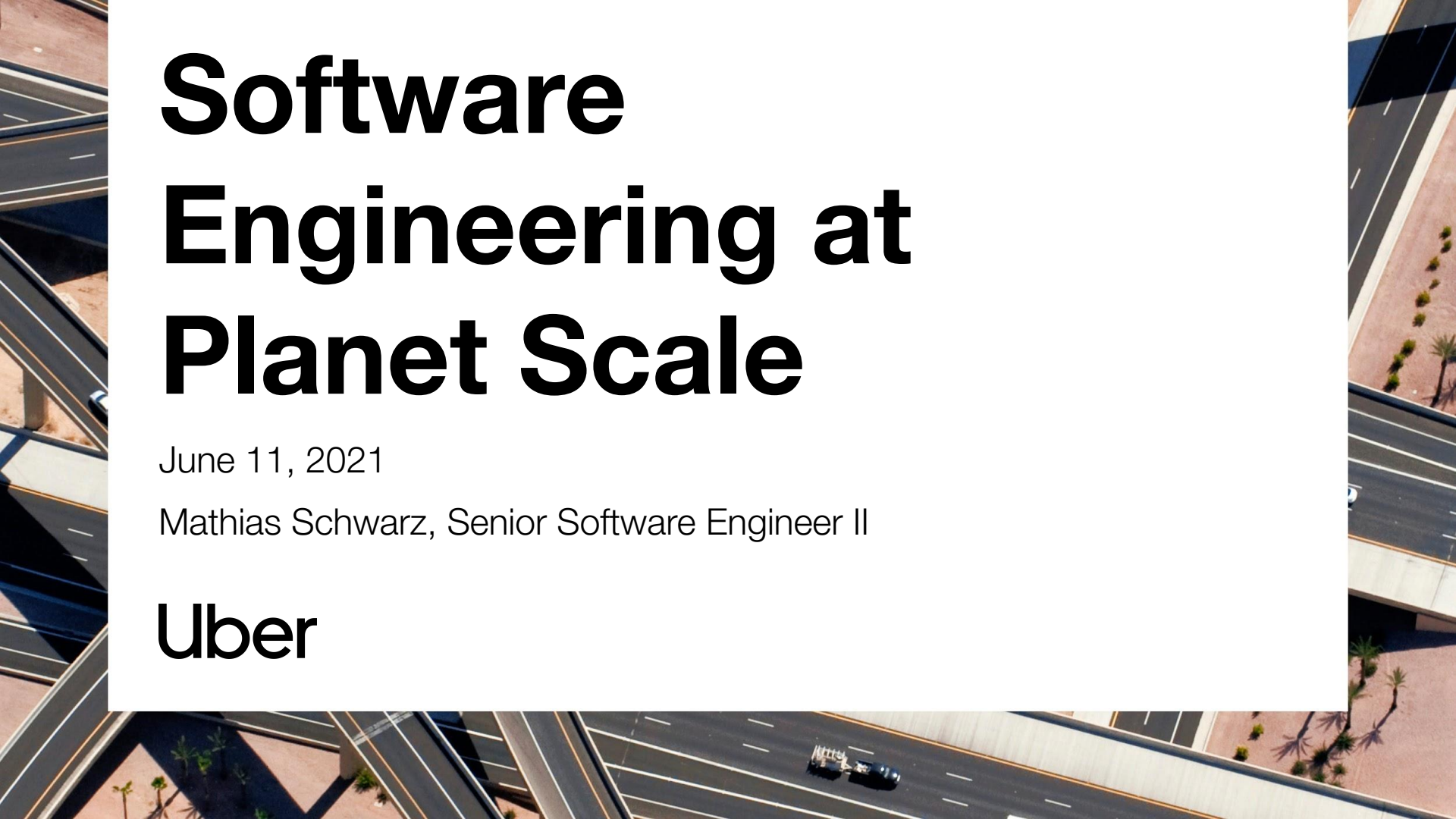
Mathias Schwarz

PhD in CS from Aarhus University

Senior Software Engineer II at Uber

Love to build infrastructure software

6 years into building stateless platforms at Uber



Software Engineering at Planet Scale

June 11, 2021

Mathias Schwarz, Senior Software Engineer II

Uber

The Uber Infrastructure team @aarhus

Products

Product Platform

Infrastructure

- Microservice deployment
- Runtime Configuration
- Storage
- Auto scaling

Scale: Business

18M

Trips happening every day,
based on Q1 2020

Scale: Infrastructure

4000

Services across several of our own data centers plus cloud zones such as AWS

Scale: Deployment

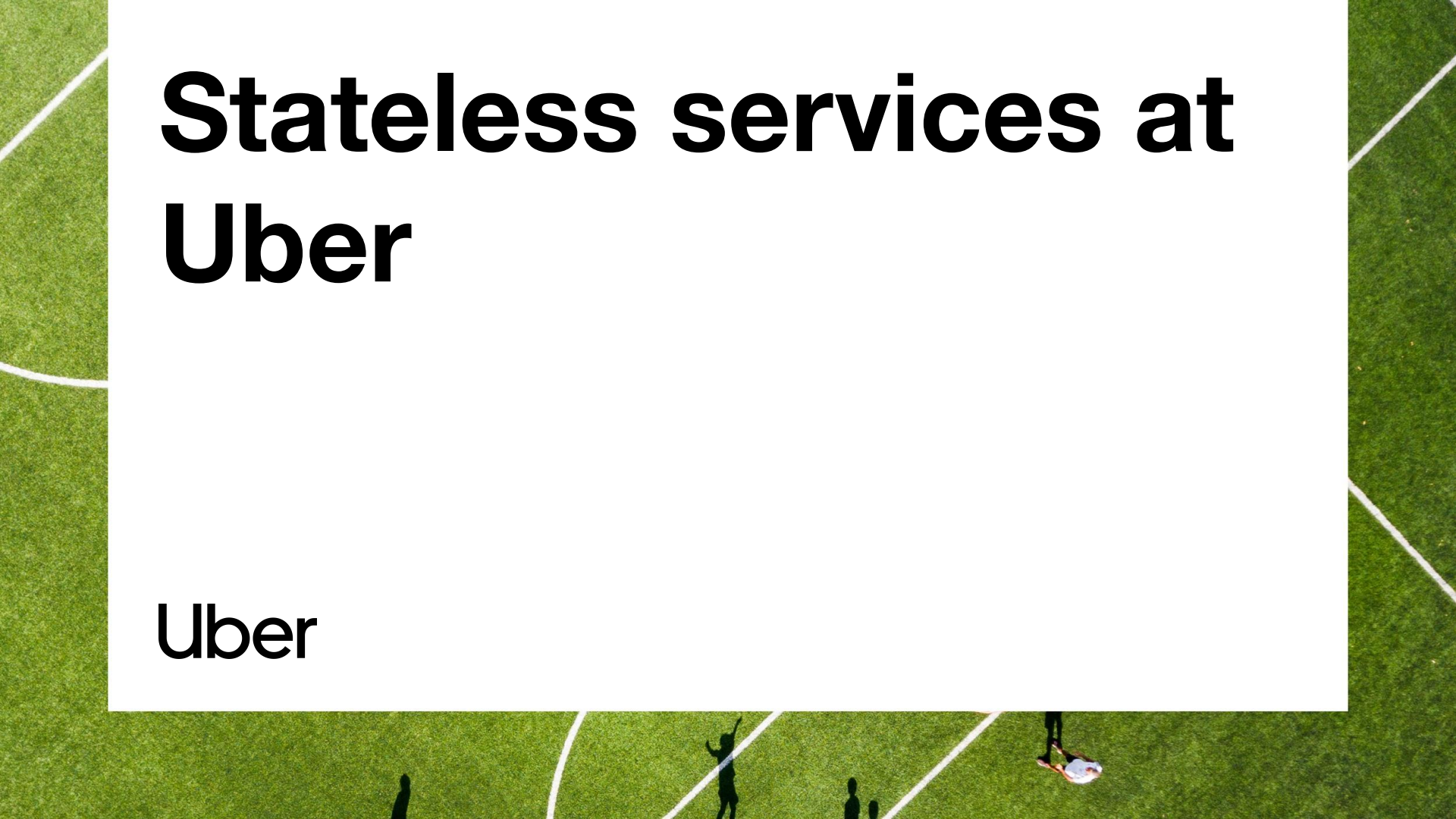
58K

Builds / week

5K

Production deploys / week

Stateless services at Uber

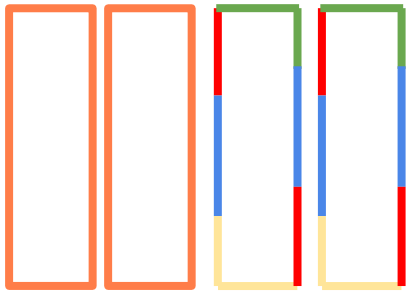


Uber

Hybrid cloud

More elastic (if you are small)

Special purpose hardware



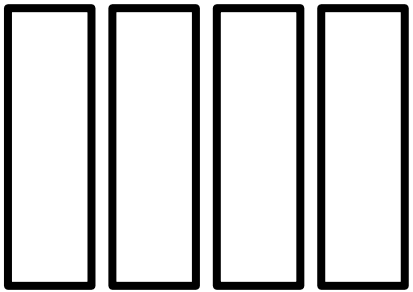
Cloud Zones / POPs

AZs with virtual Hardware

+

Cheaper

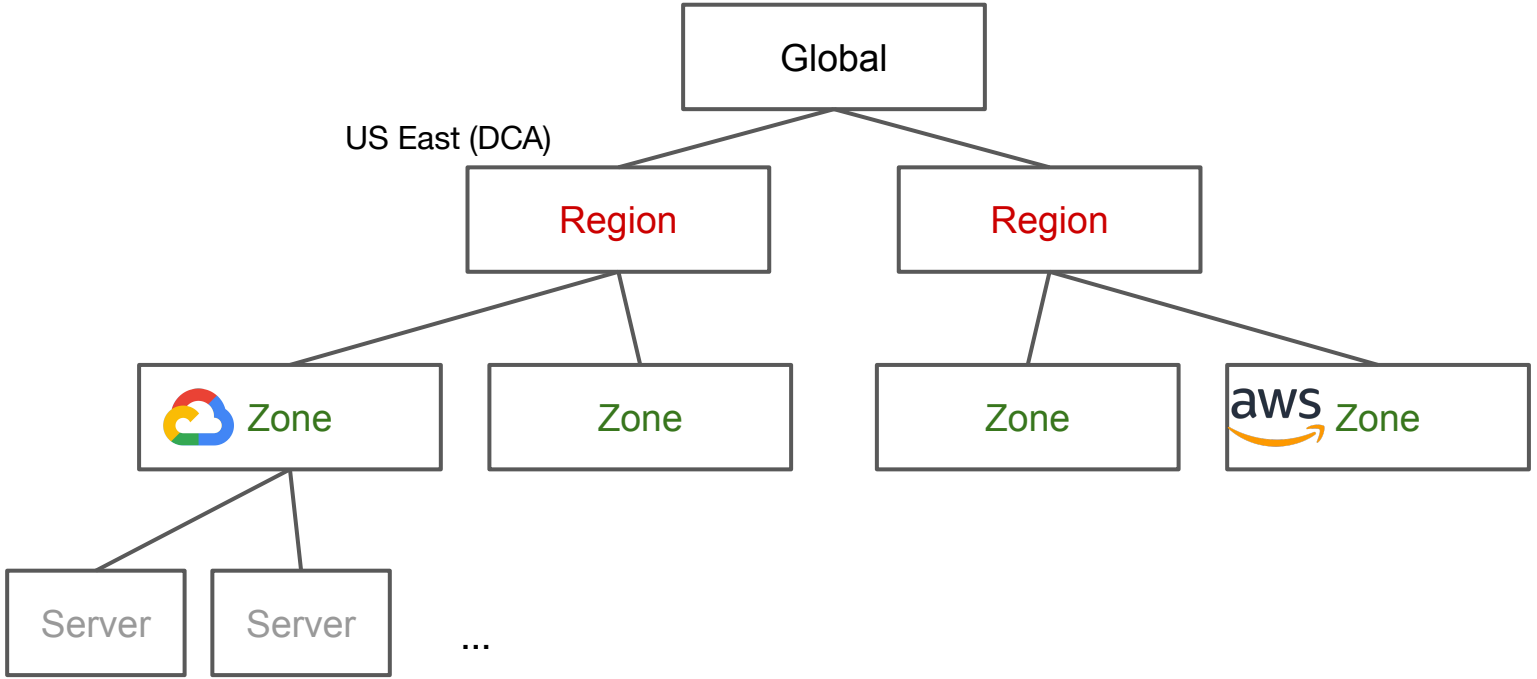
Few dependencies



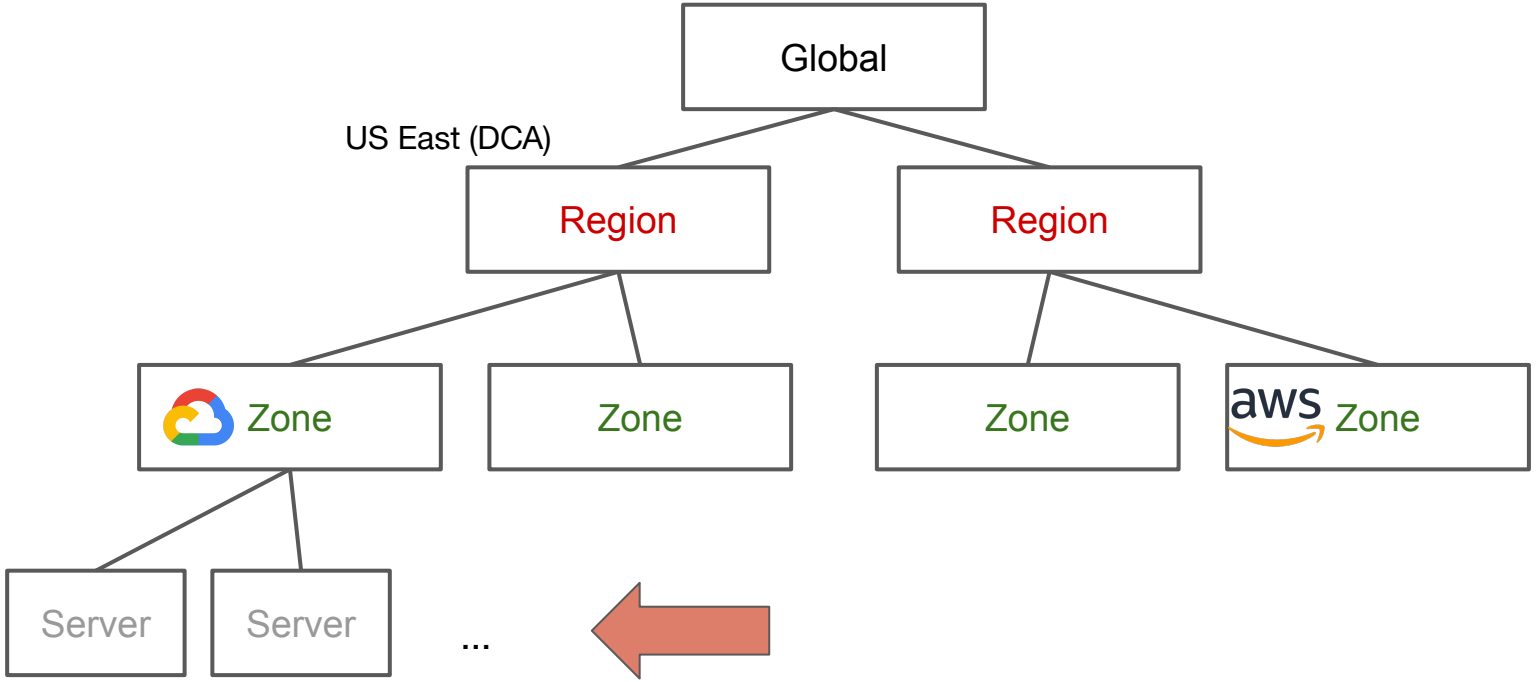
Private / On Prem

Core AZs with ODM Hardware

Regions and availability zones



Early days: Unstructured deploys



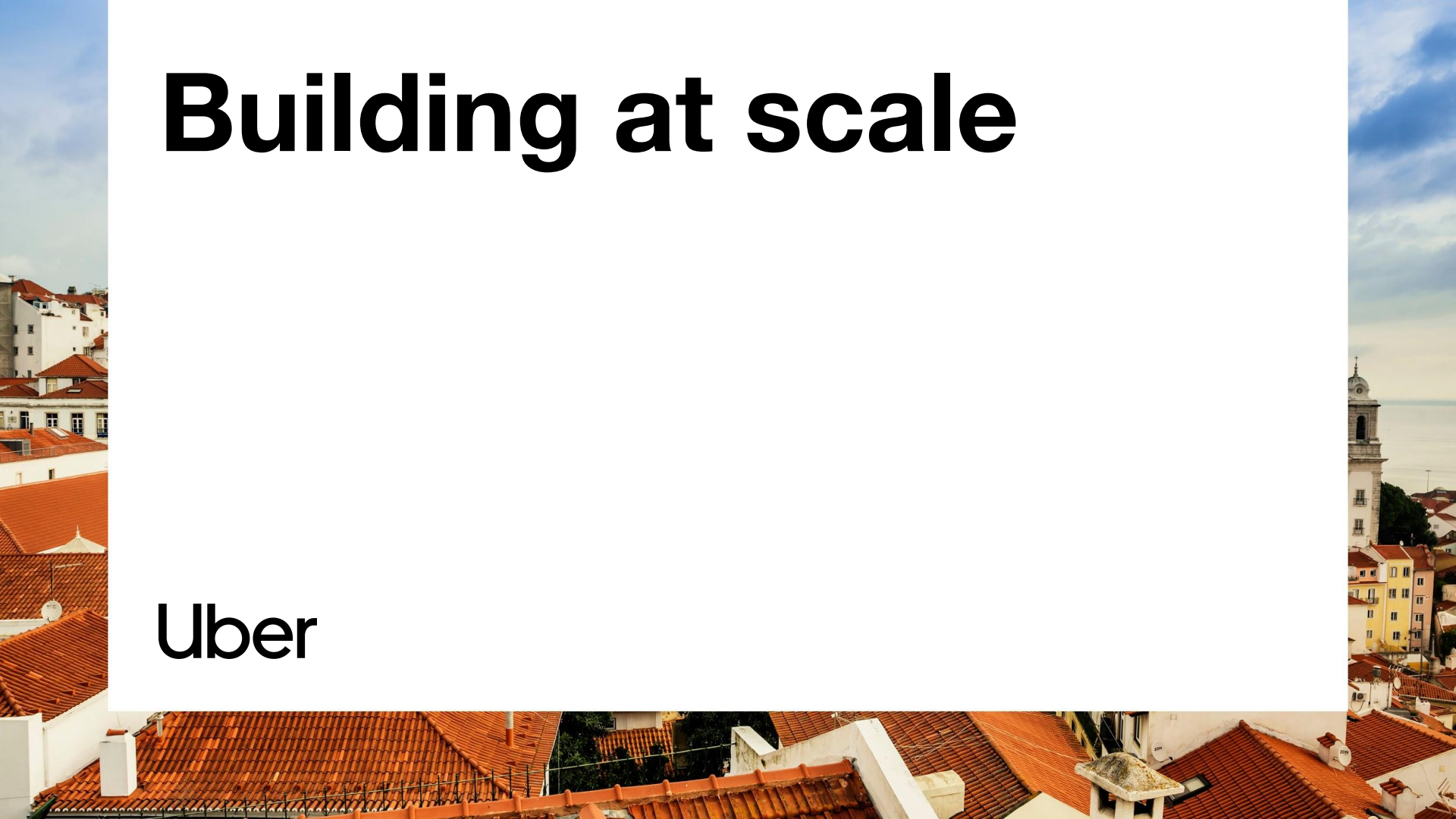
Important deploy system features

Consistent builds

Zero downtime

Outage prevention

Building at scale



Uber

Code repository

Go & Java

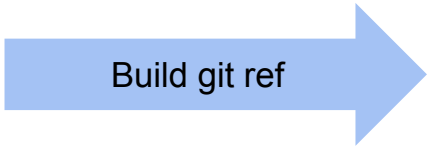
The preferred backend language

Bazel

Monorepo management

https://eng.uber.com/go-monorepo-bazel/

Build a docker image



Makisu

+ uBuild

Builds docker images

https://github.com/uber/makisu
https://eng.uber.com/makisu/

Publish docker image

Makisu
+uBuild

Upload

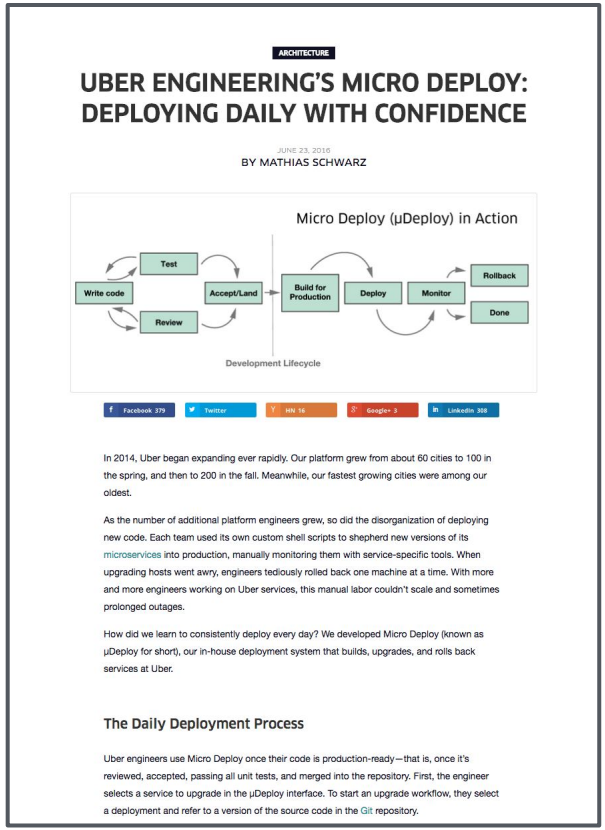
Kraken

Distributed P2P Docker Registry

https://github.com/uber/kraken ,
https://eng.uber.com/introducing-kraken/

Structured deploys with µDeploy

Uber



ARCHITECTURE

UBER ENGINEERING'S MICRO DEPLOY: DEPLOYING DAILY WITH CONFIDENCE

JUNE 23, 2016

BY MATHIAS SCHWARZ

In 2014, Uber began expanding ever rapidly. Our platform grew from about 60 cities to 100 in the spring, and then to 200 in the fall. Meanwhile, our fastest growing cities were among our oldest.

As the number of additional platform engineers grew, so did the disorganization of deploying new code. Each team used its own custom shell scripts to shepherd new versions of its microservices into production, manually monitoring them with service-specific tools. When upgrading hosts went awry, engineers tediously rolled back one machine at a time. With more and more engineers working on Uber services, this manual labor couldn't scale and sometimes prolonged outages.

How did we learn to consistently deploy every day? We developed Micro Deploy (known as µDeploy for short), our in-house deployment system that builds, upgrades, and rolls back services at Uber.

The Daily Deployment Process

Uber engineers use Micro Deploy once their code is production-ready—that is, once it's reviewed, accepted, passing all unit tests, and merged into the repository. First, the engineer selects a service to upgrade in the µDeploy interface. To start an upgrade workflow, they select a deployment and refer to a version of the source code in the Git repository.

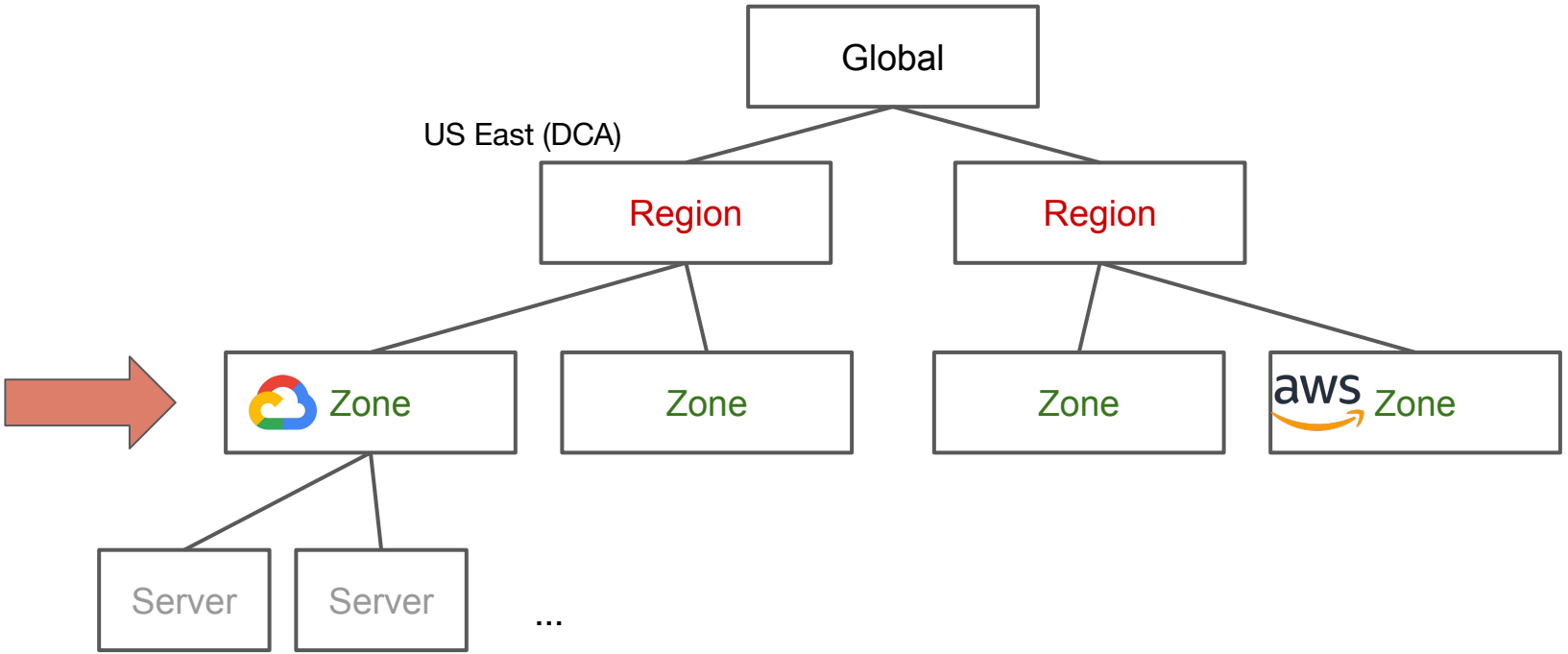
Rollout to clusters, not servers

Scheduler+Manager for compute clusters

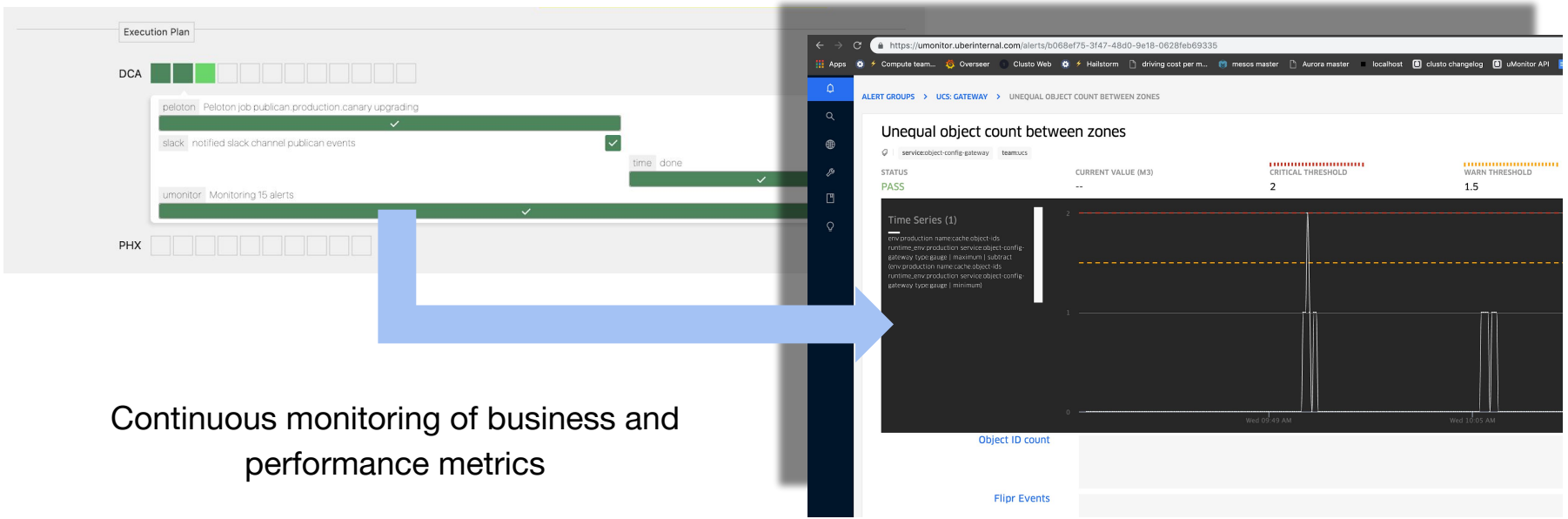
https://github.com/uber/peloton

https://eng.uber.com/resource-scheduler-cluster-management-peloton/

Deploy into cluster in zones



Safety: Monitoring metrics

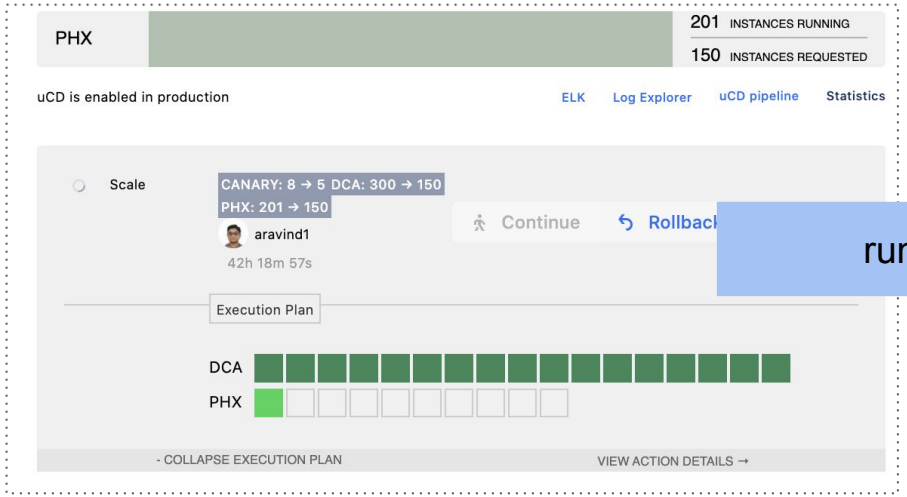


Continuous monitoring of business and performance metrics

uMonitor

time-series based rollback triggers

Safety: Whitebox integration



run

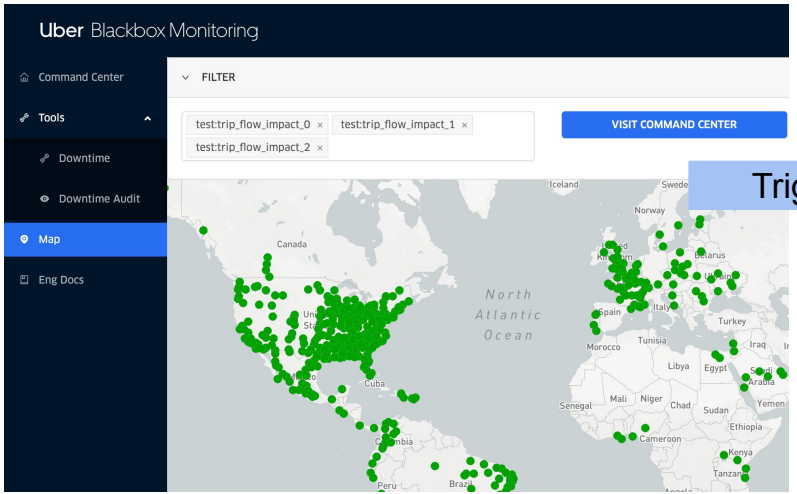
Explicit part of the rollout plan

Hailstorm

Load tests

Whitebox integration tests

Safety: Continuous blackbox



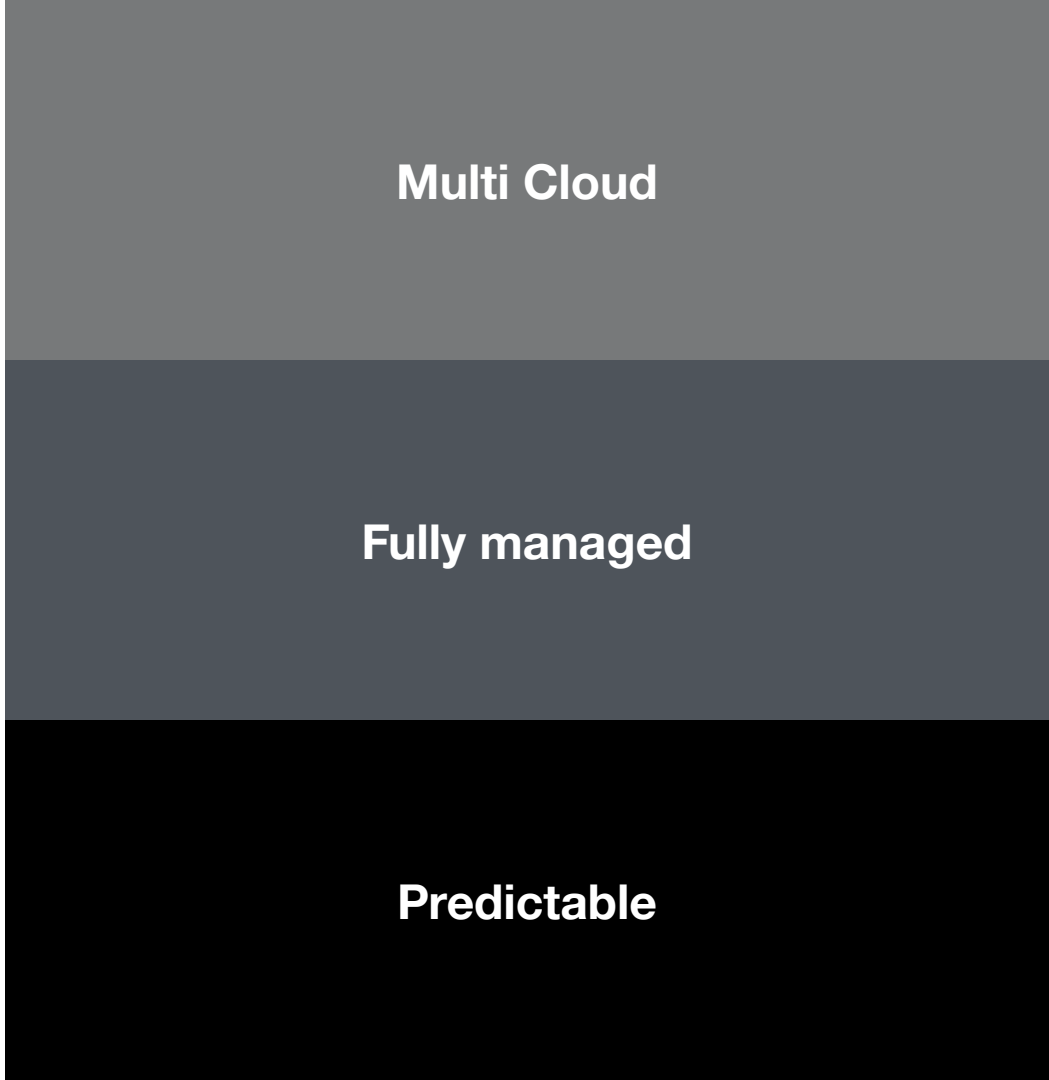
Trigger page

Engineers decide on rolling back

Blackbox testing/monitoring

Virtual trips in production

Efficiency at scale



Alternatives

”Deploy across multiple cloud providers including AWS EC2, Kubernetes, Google Compute Engine, Google Kubernetes Engine, Google App Engine, Microsoft Azure, Openstack, Cloud Foundry, and Oracle Cloud Infrastructure, with DC/OS coming soon”

Anthos

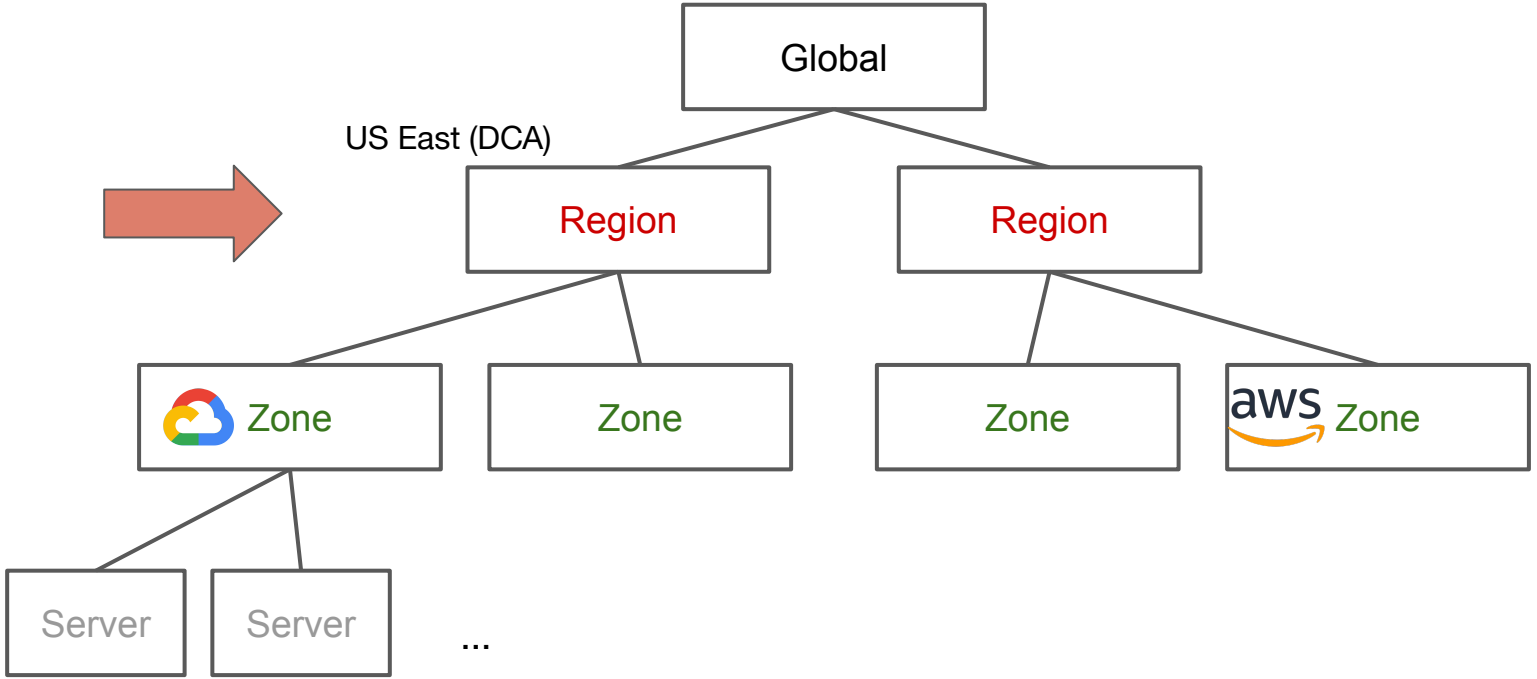
GitLab

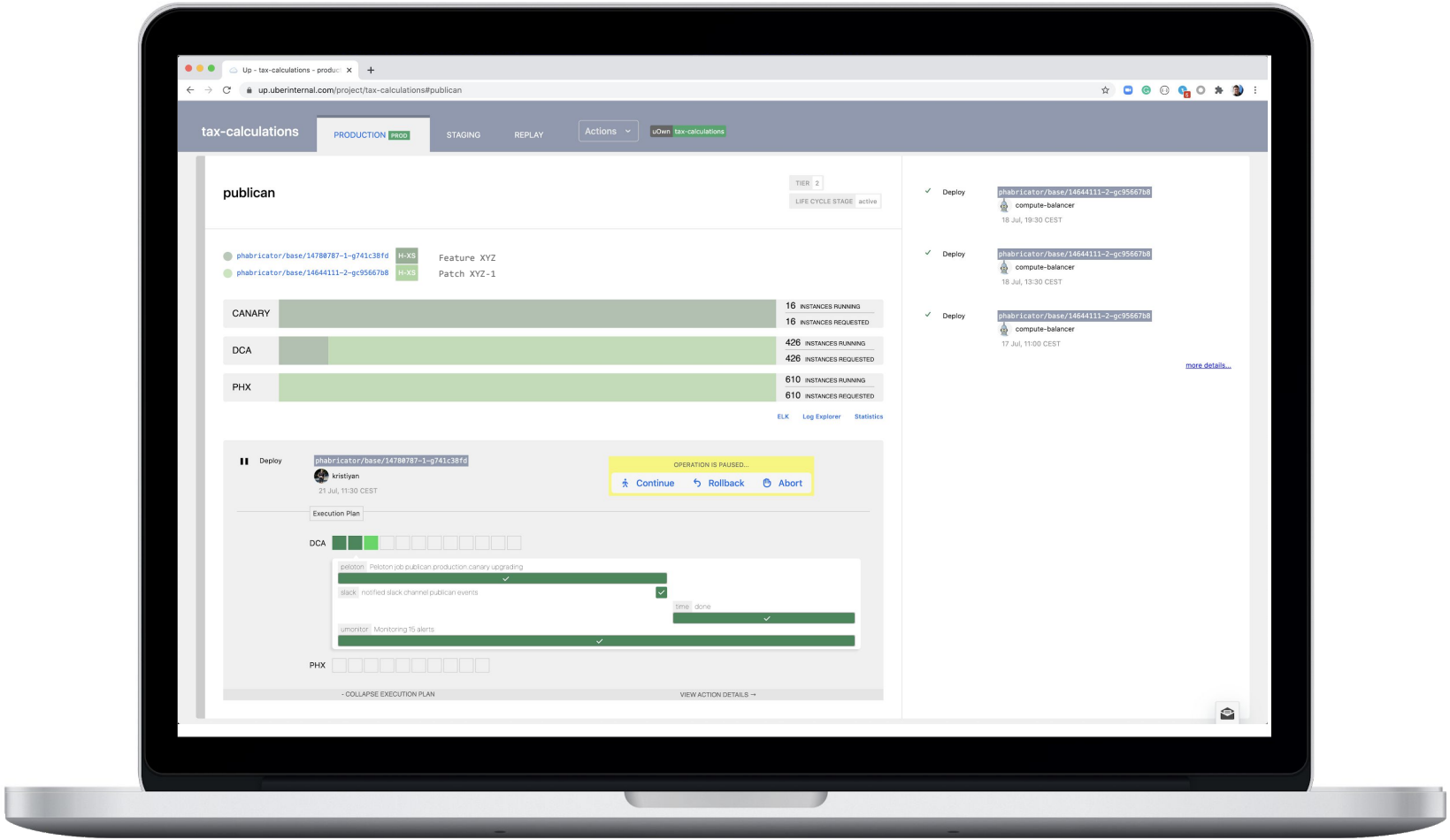
https://spinnaker.io/

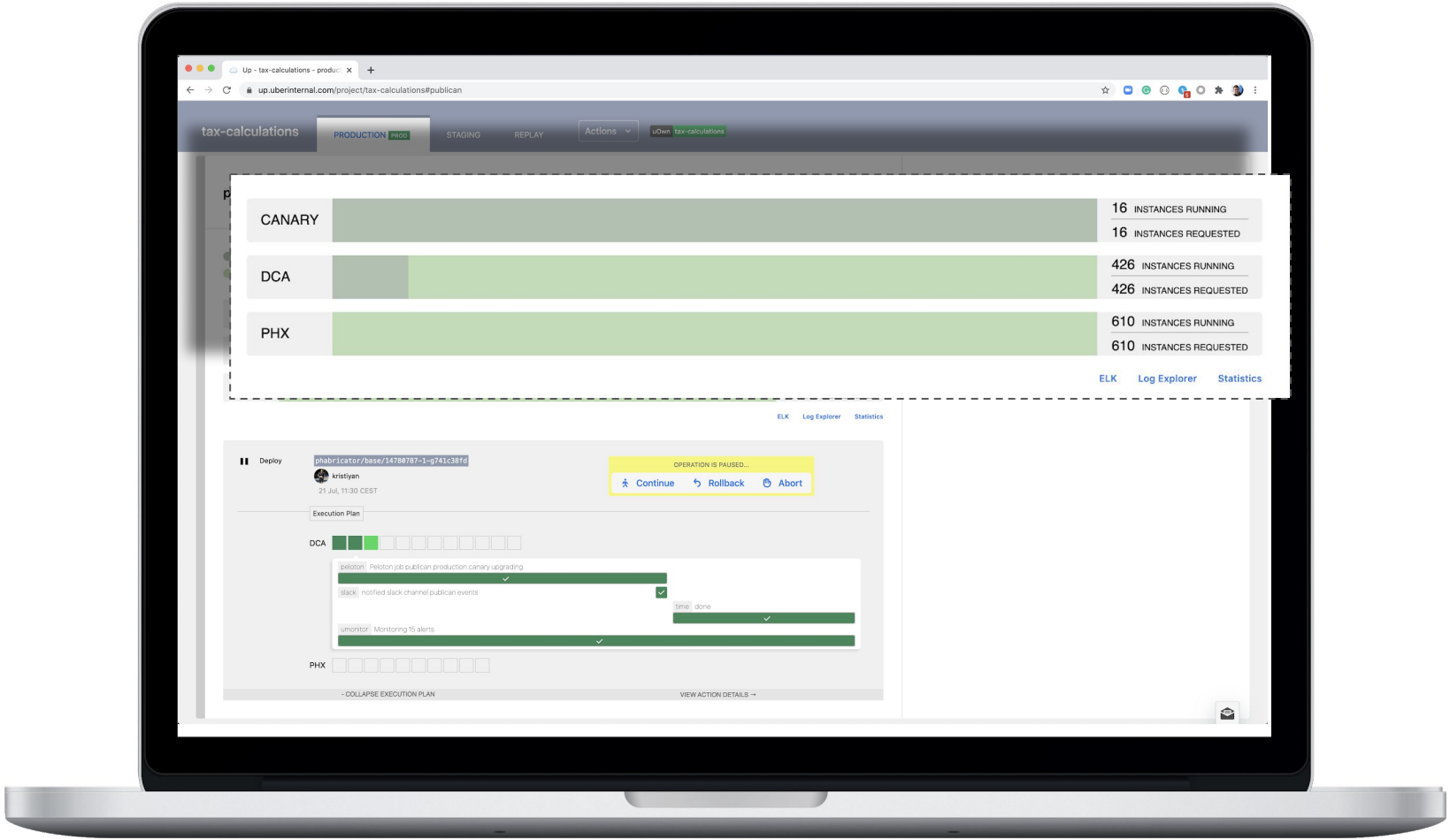
Efficient infrastructure with Up

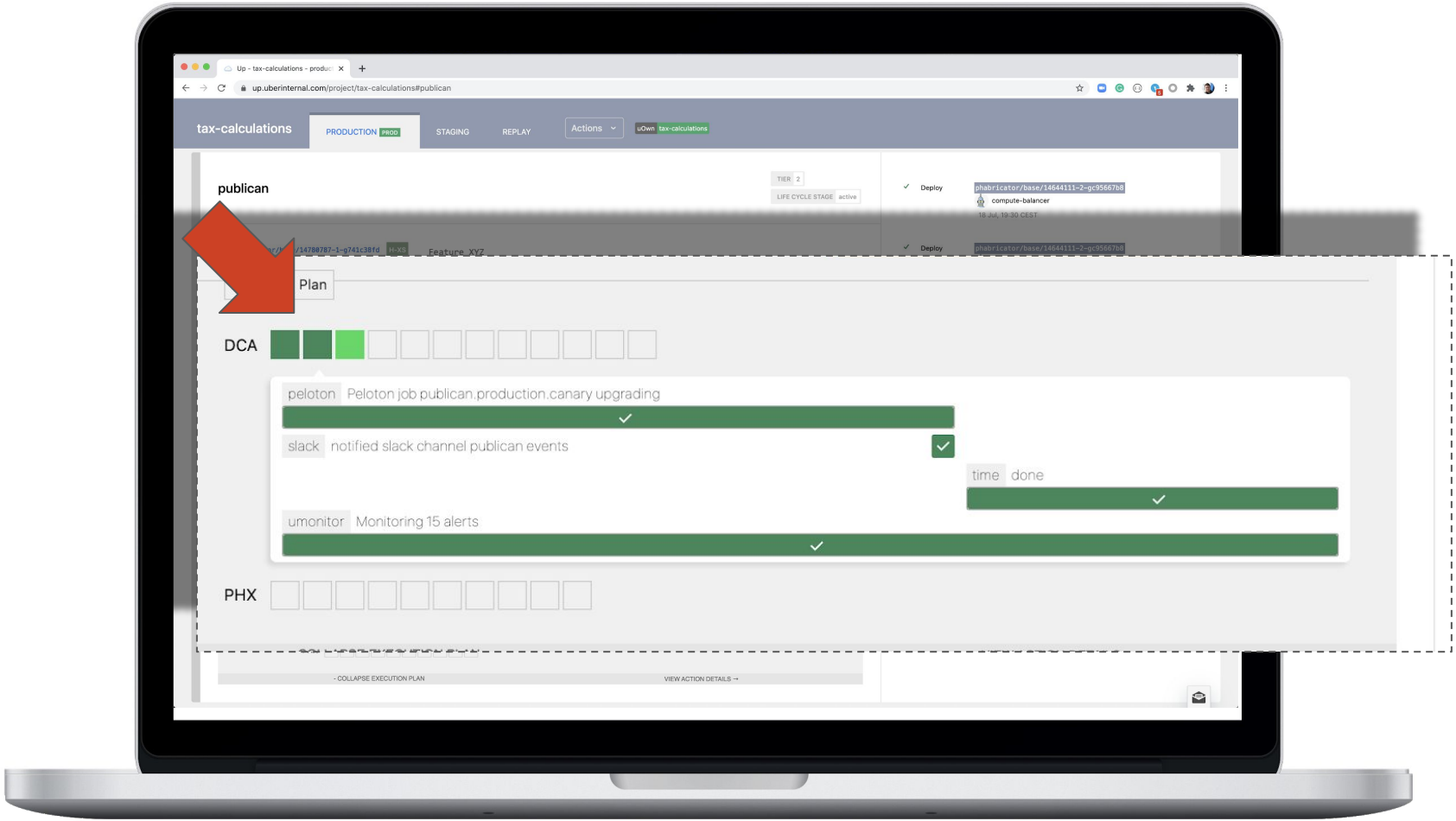
Uber

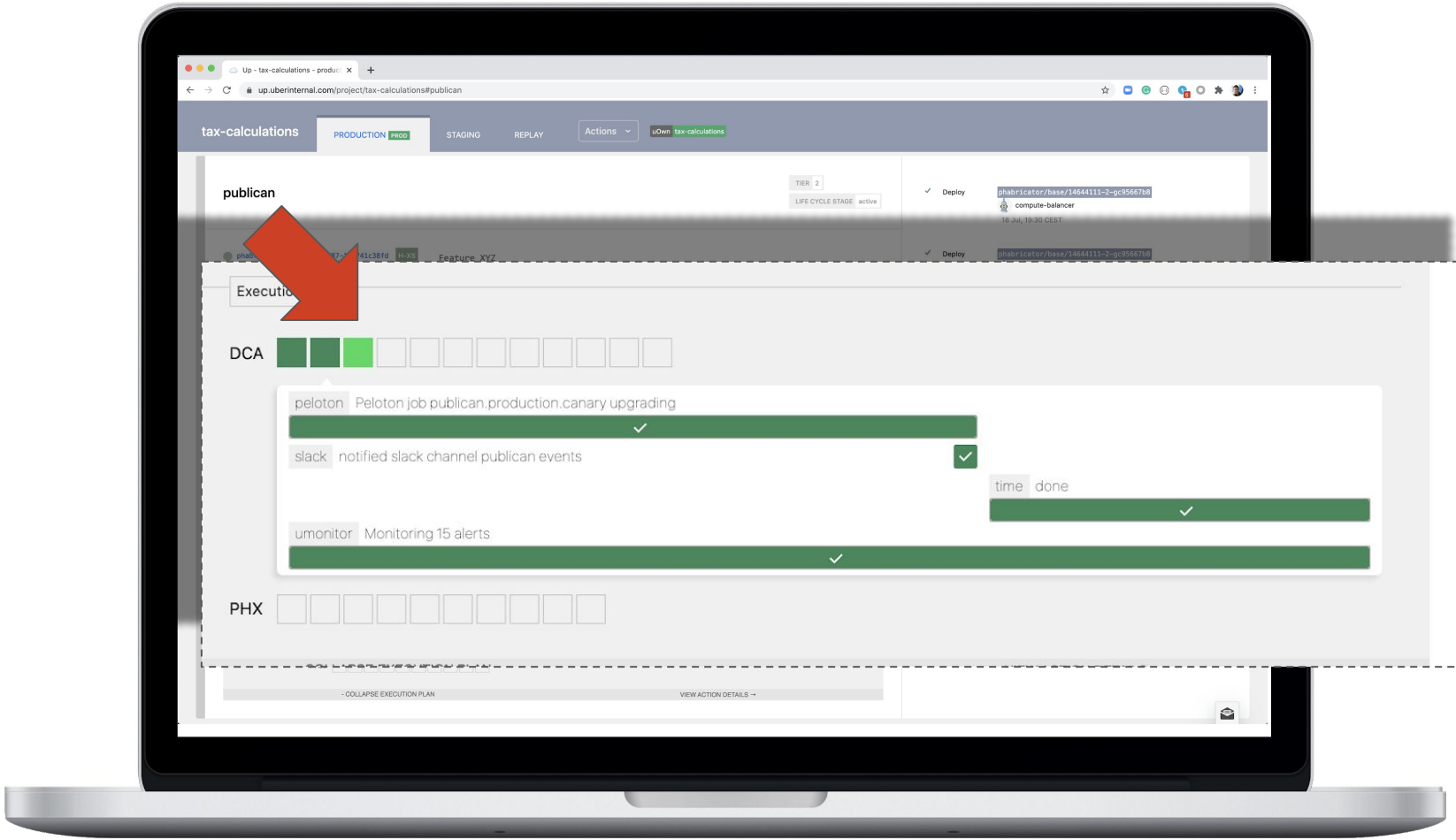
Deploy into regions

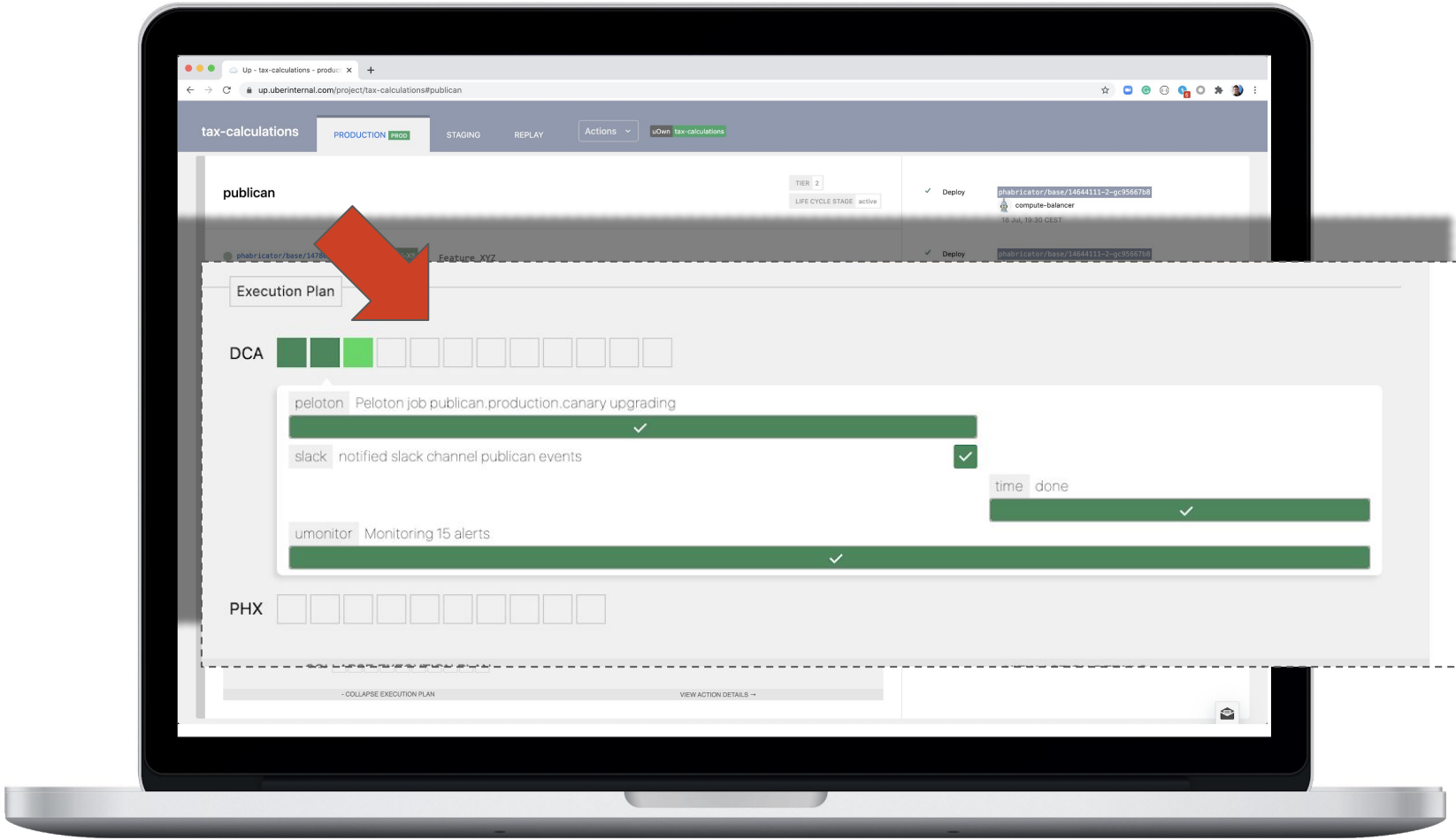



up.uberinternal.com/project/tax-calculations#publican
tax-calculations
PRODUCTION PROD
STAGING
REPLAY
Actions
uOwn tax-calculations
publican
TIER 2
LIFE CYCLE STAGE active
phabricator/base/14700787-1-g741c38fd H-XS Feature XYZ
phabricator/base/14644111-2-gc95667b8 H-XS Patch XYZ-1
CANARY
16 INSTANCES RUNNING
16 INSTANCES REQUESTED
DCA
426 INSTANCES RUNNING
426 INSTANCES REQUESTED
PHX
610 INSTANCES RUNNING
610 INSTANCES REQUESTED
ELK
Log Explorer
Statistics
Deploy
phabricator/base/14700787-1-g741c38fd
kristiyan
21 Jul, 11:30 CEST
OPERATION IS PAUSED...
Continue
Rollback
Abort
Execution Plan
DCA
peloton Peloton job publican.production.canary.upgrading
slack notified slack channel publican events
time done
umonitor Monitoring 15 alerts
PHX
- COLLAPSE EXECUTION PLAN
VIEW ACTION DETAILS →
Deploy
phabricator/base/14644111-2-gc95667b8
compute-balancer
18 Jul, 19:30 CEST
Deploy
phabricator/base/14644111-2-gc95667b8
compute-balancer
18 Jul, 13:30 CEST
Deploy
phabricator/base/14644111-2-gc95667b8
compute-balancer
17 Jul, 11:00 CEST
more details...


up.uberinternal.com/project/tax-calculations#publican
tax-calculations
PRODUCTION PROD
STAGING
REPLAY
Actions
uOwn tax-calculations
CANARY
16 INSTANCES RUNNING
16 INSTANCES REQUESTED
DCA
426 INSTANCES RUNNING
426 INSTANCES REQUESTED
PHX
610 INSTANCES RUNNING
610 INSTANCES REQUESTED
ELK
Log Explorer
Statistics
Deploy
phabricator/base/14700787-1-g741c38fd
kristiyan
21 Jul, 11:30 CEST
OPERATION IS PAUSED...
Continue
Rollback
Abort
Execution Plan
DCA
peloton Peloton job publican.production.canary upgrading
slack notified slack channel publican events
time done
umonitor Monitoring 15 alerts
PHX
- COLLAPSE EXECUTION PLAN
VIEW ACTION DETAILS →


up.uberinternal.com/project/tax-calculations#publican
tax-calculations
PRODUCTION PROD
STAGING
REPLAY
Actions
uOwn tax-calculations
publican
TIER 2
LIFE CYCLE STAGE active
Deploy
phabricator/base/14644111-2-gc95667b8
compute-balancer
18 Jul, 19:30 CEST
Deploy
phabricator/base/14644111-2-gc95667b8
Feature_XYZ
Plan
DCA
peloton Peloton job publican.production.canary upgrading
slack notified slack channel publican events
time done
umonitor Monitoring 15 alerts
PHX
- COLLAPSE EXECUTION PLAN
VIEW ACTION DETAILS →

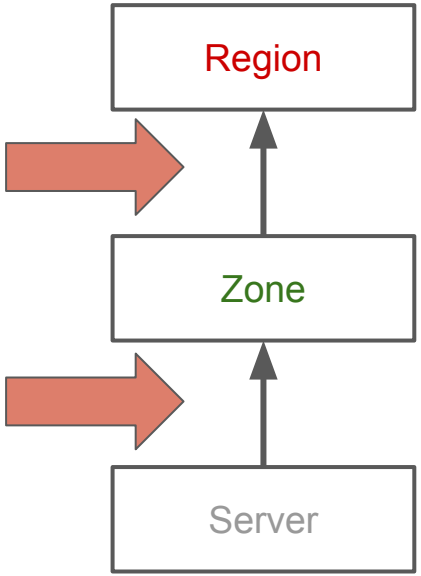

up.uberinternal.com/project/tax-calculations#publican
tax-calculations
PRODUCTION PROD
STAGING
REPLAY
Actions
uOwn tax-calculations
publican
TIER 2
LIFE CYCLE STAGE active
Deploy
phabricator/base/14644111-2-gc95667b8
compute-balancer
18 Jul, 19:30 CEST
Deploy
phabricator/base/14644111-2-gc95667b8
Feature_XYZ
Executio
DCA
peloton Peloton job publican.production.canary upgrading
slack notified slack channel publican events
time done
umonitor Monitoring 15 alerts
PHX
- COLLAPSE EXECUTION PLAN
VIEW ACTION DETAILS →


up.uberinternal.com/project/tax-calculations#publican
tax-calculations
PRODUCTION PROD
STAGING
REPLAY
Actions
uOwn tax-calculations
publican
TIER 2
LIFE CYCLE STAGE active
Deploy
phabricator/base/14644111-2-gc95667b8
compute-balancer
18 Jul, 19:30 CEST
Deploy
phabricator/base/14644111-2-gc95667b8
Feature_XYZ
Execution Plan
DCA
peloton Peloton job publican.production.canary upgrading
slack notified slack channel publican events
time done
umonitor Monitoring 15 alerts
PHX
- COLLAPSE EXECUTION PLAN
VIEW ACTION DETAILS →

Challenges in abstraction

Dependency on low latency

Dependency on host data



Avoiding manual migrations

Uber

Plan for migrations

- More productive teams
- More reliable / available
- Faster
- Lower cost of capacity

- Cloud <-> On-prem zones
- Zones come and go
- Cluster rolls change
- New cluster technologies
- New container runtimes
- New base images
- ….

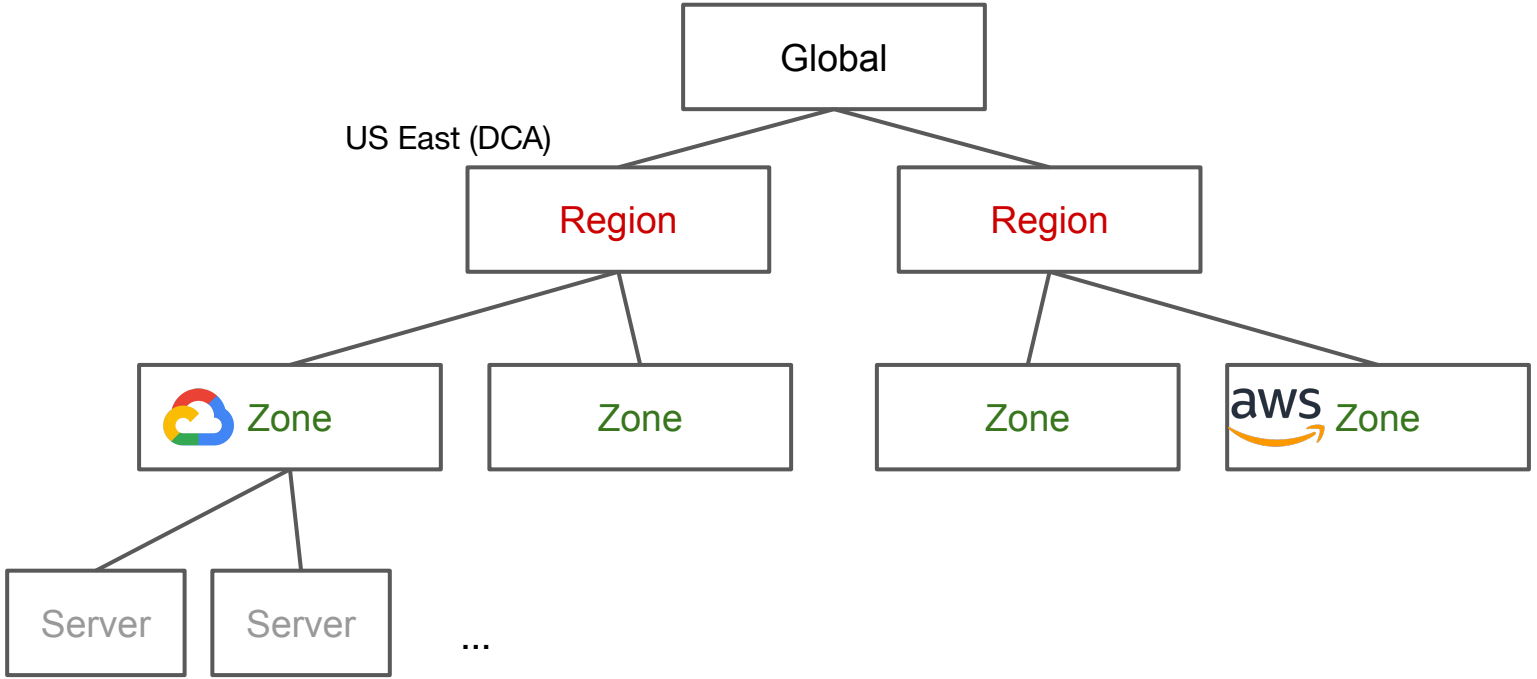
Goals

A continuous need for improving efficiency and reducing the cost of running the company

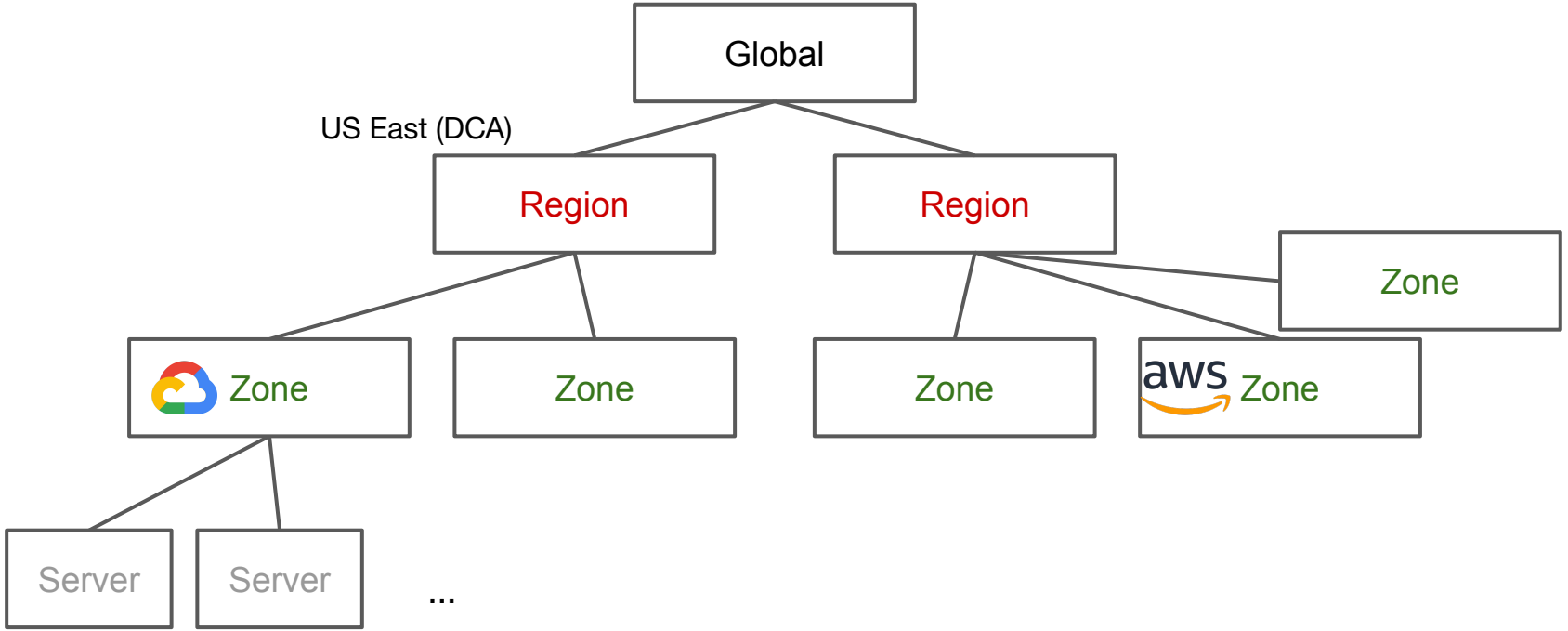
Infra changes

Must **not** have a negative impact on the goals

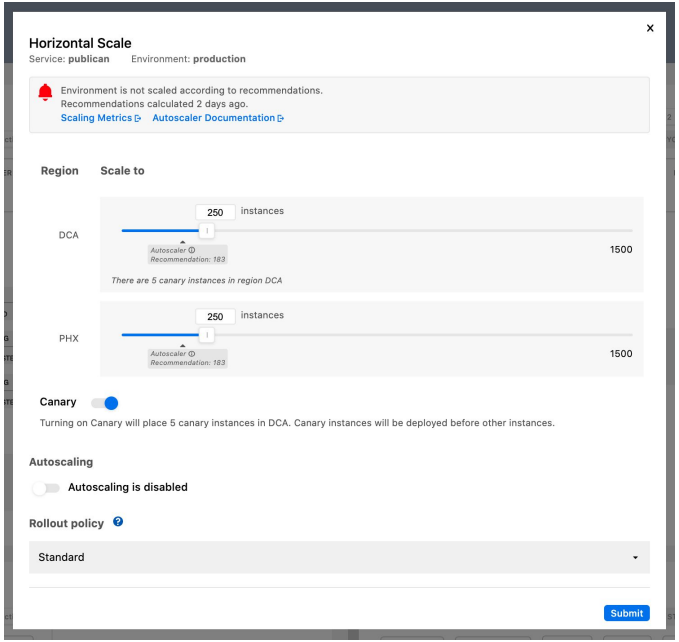
Let’s add a new zone

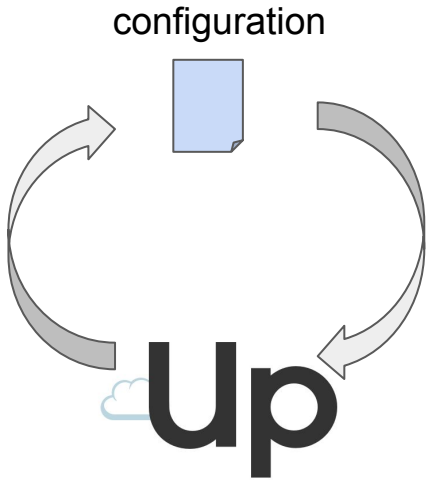


Let’s add a new zone



Declarative Configuration



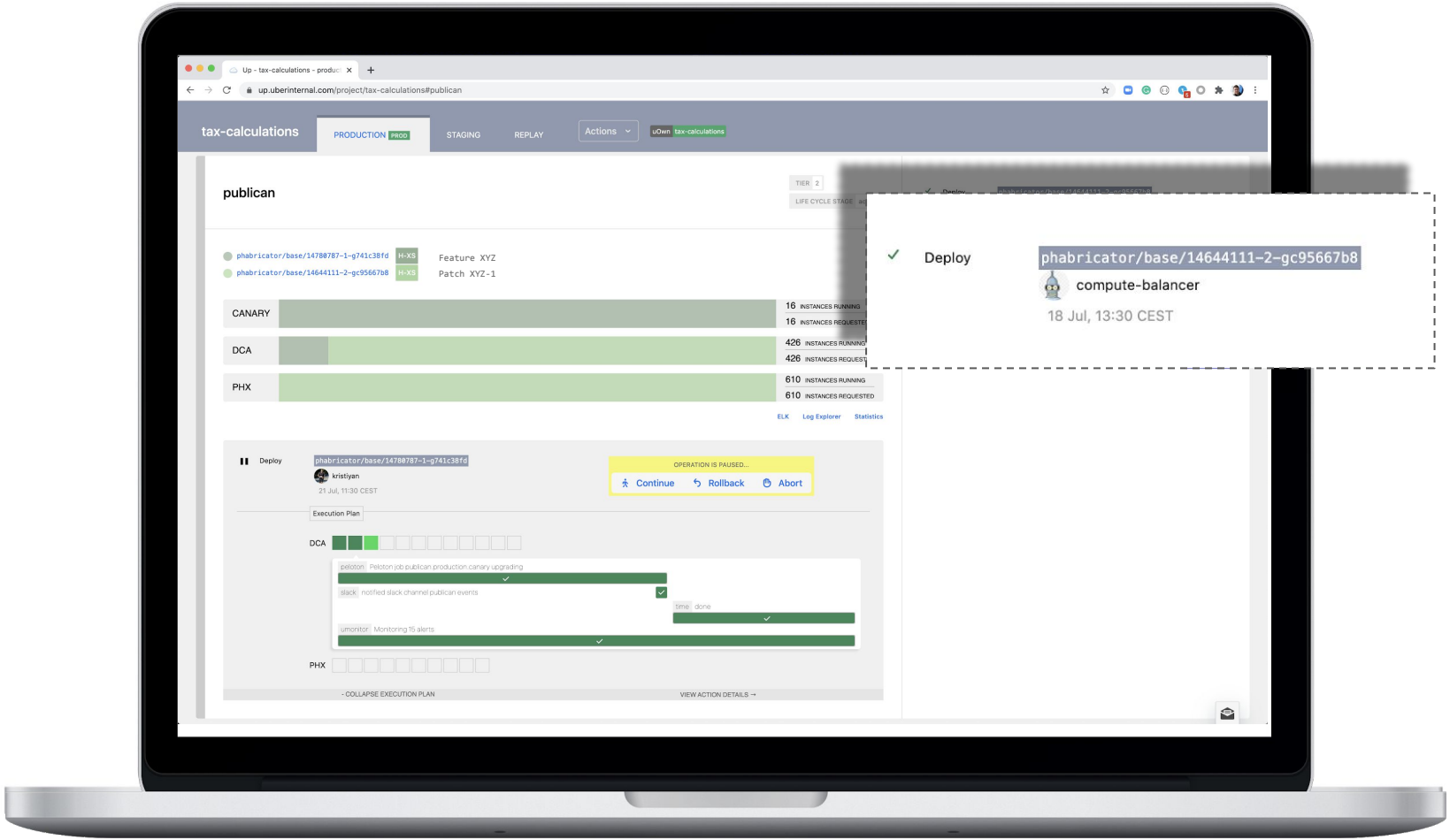


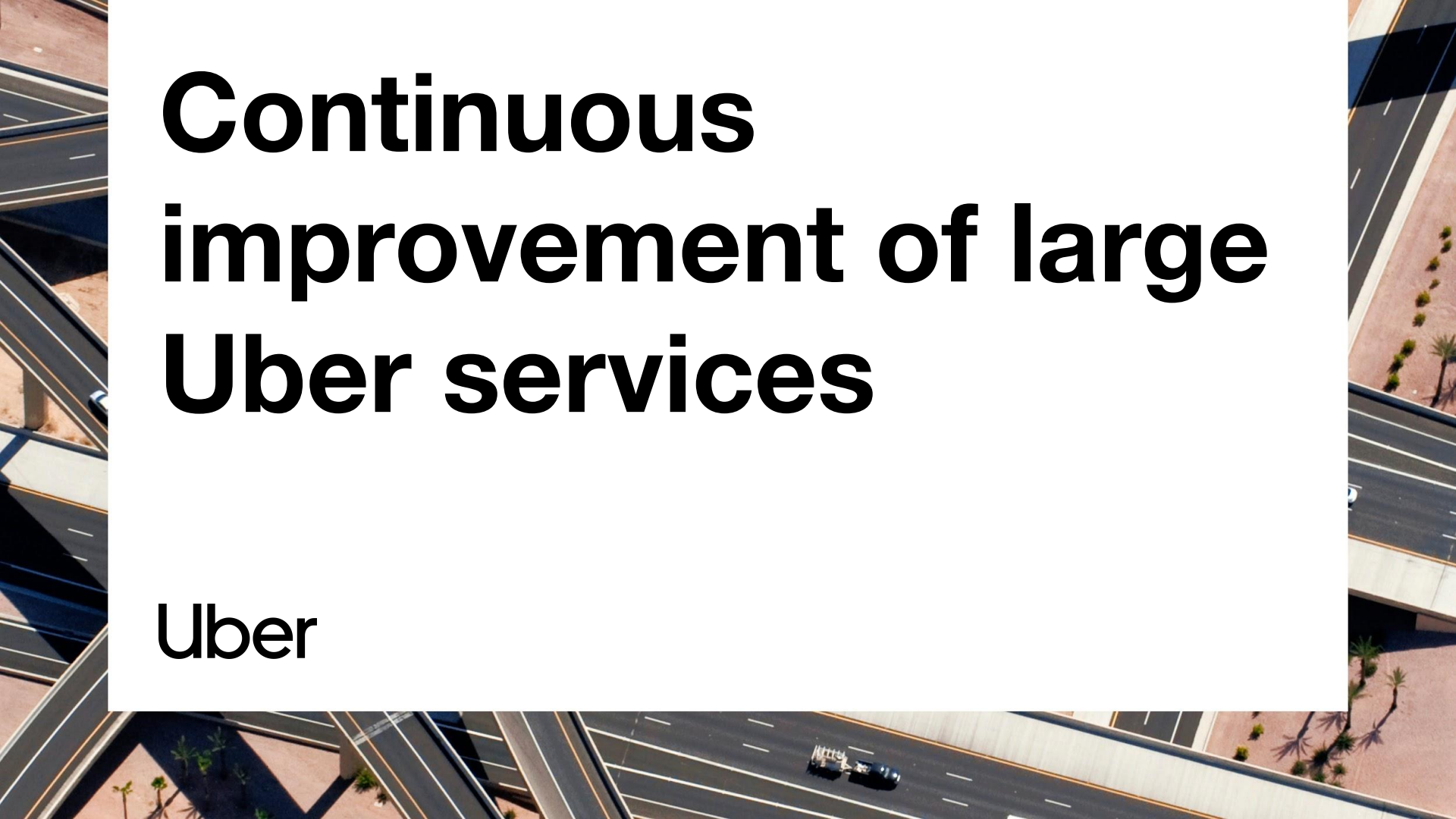
Declarative configurations

- Region placement
- Dependencies

Continuous evaluation

Optimal application of the configuration


up.uberinternal.com/project/tax-calculations#publican
tax-calculations
PRODUCTION PROD
STAGING
REPLAY
Actions
uOwn tax-calculations
publican
phabricator/base/14700787-1-g741c38fd H-XS Feature XYZ
phabricator/base/14644111-2-gc95667b8 H-XS Patch XYZ-1
CANARY
16 INSTANCES RUNNING
16 INSTANCES REQUESTED
DCA
426 INSTANCES RUNNING
426 INSTANCES REQUESTED
PHX
610 INSTANCES RUNNING
610 INSTANCES REQUESTED
ELK
Log Explorer
Statistics
Deploy
phabricator/base/14700787-1-g741c38fd
kristiyan
21 Jul, 11:30 CEST
OPERATION IS PAUSED...
Continue
Rollback
Abort
Execution Plan
DCA
peloton Peloton job publican.production.canary.upgrading
slack notified slack channel publican events
time done
umonitor Monitoring 15 alerts
PHX
- COLLAPSE EXECUTION PLAN
VIEW ACTION DETAILS →
Deploy
phabricator/base/14644111-2-gc95667b8
compute-balancer
18 Jul, 13:30 CEST



Continuous improvement of large Uber services

Uber

Metrics

Prometheus on M3DB

Metrics ingestion platform

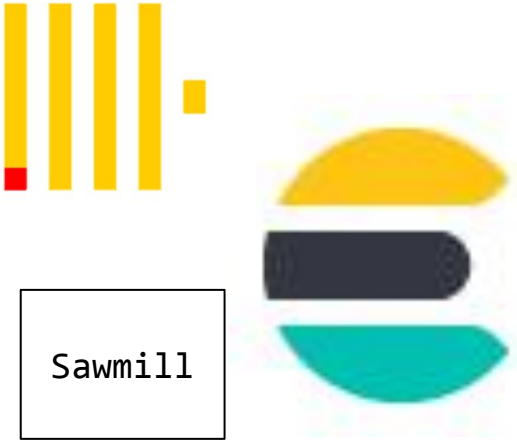
Feeds

Grafana

Graphs and dashboards

Log indexing and aggregation

Healthline

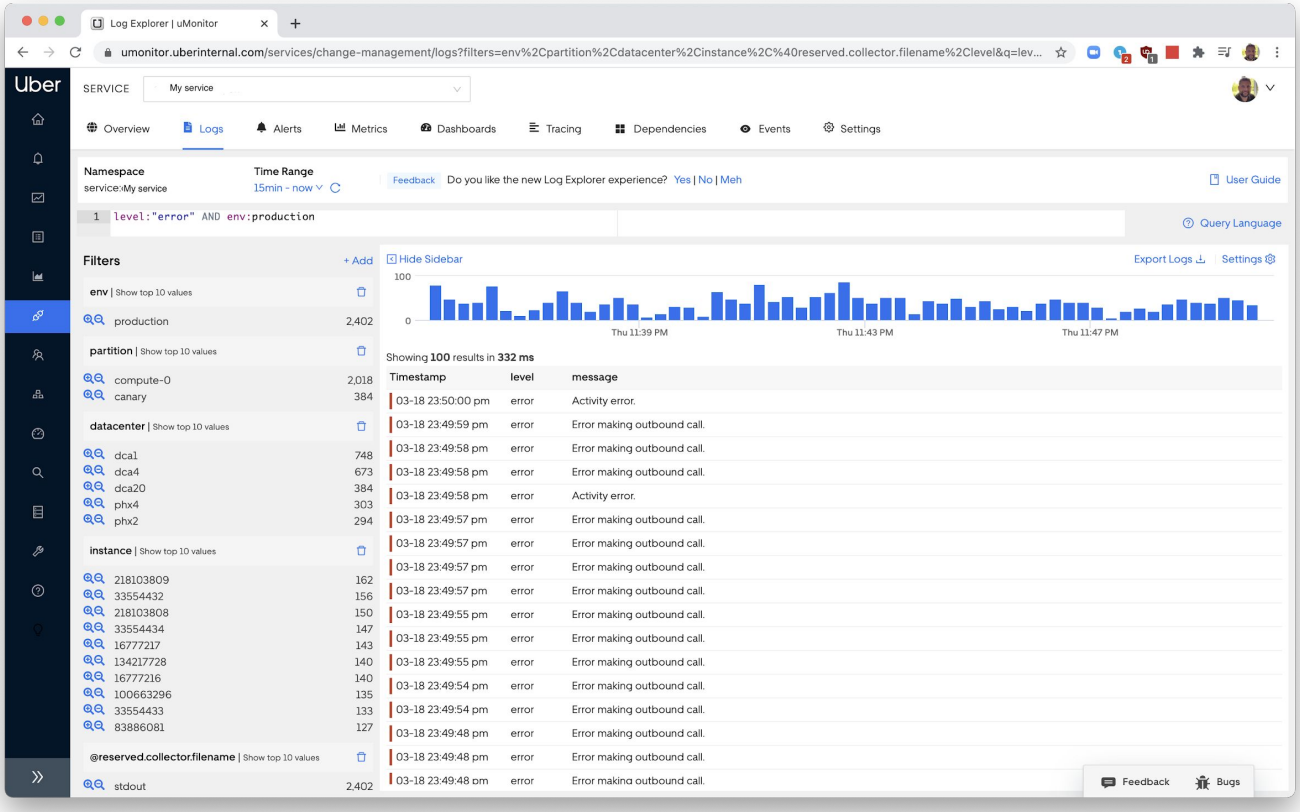


Healthline

Error log aggregation

Elasticsearch, Clickhouse, Sawmill

Log Aggregation/Queries



Distributed tracing

Jaeger

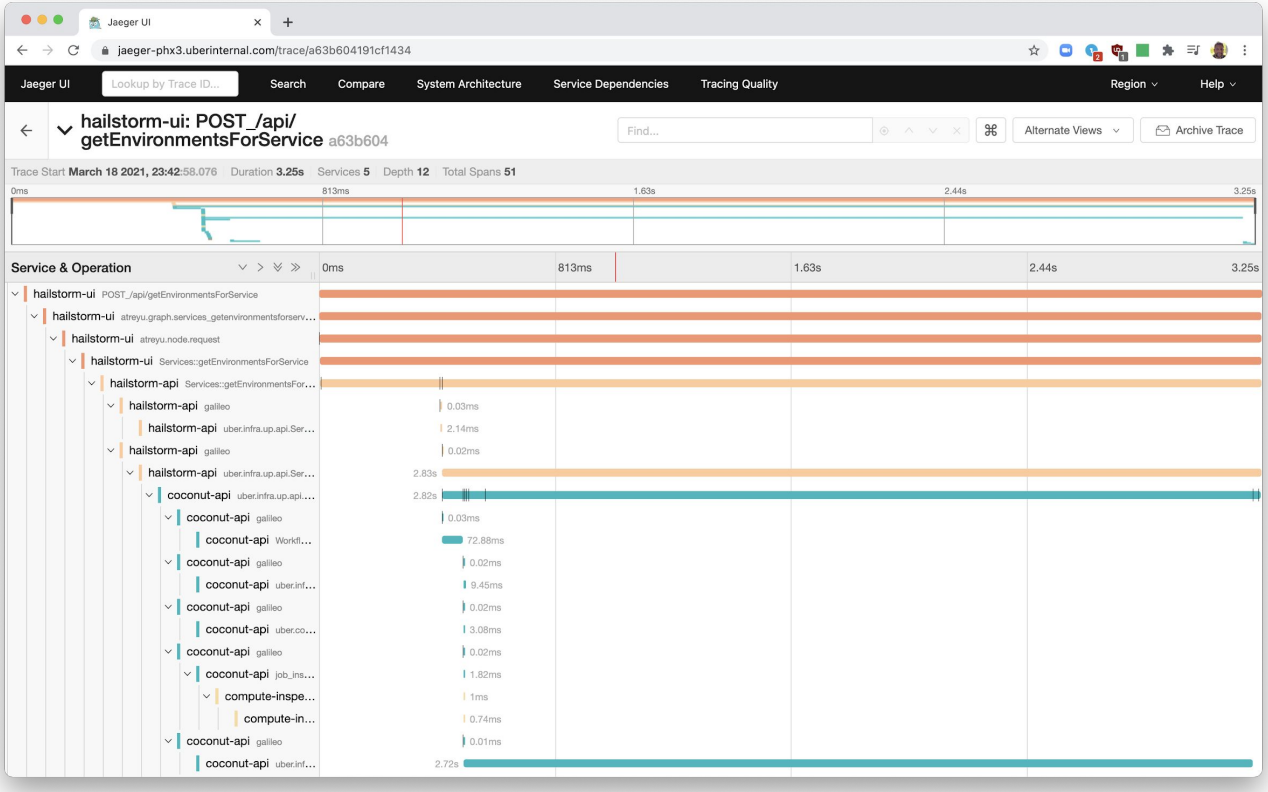
Distributed tracing for all services

https://github.com/jaegertracing/jaeger
https://opentelemetry.io/

Snapshot of callgraph

All 4K microservices and how they interact

Distributed tracing

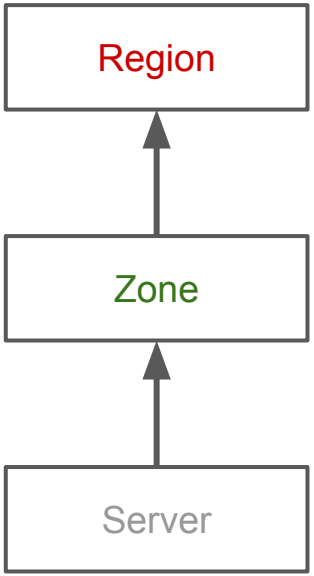


Summary

You can scale every aspect of software engineering

You can safely roll out 5000 times to production per week

Automation + abstractions are key to avoid constant manual migrations



Questions?

goto;

Don’t forget to
vote for this session
in the **GOTO Guide app**

Backup slides