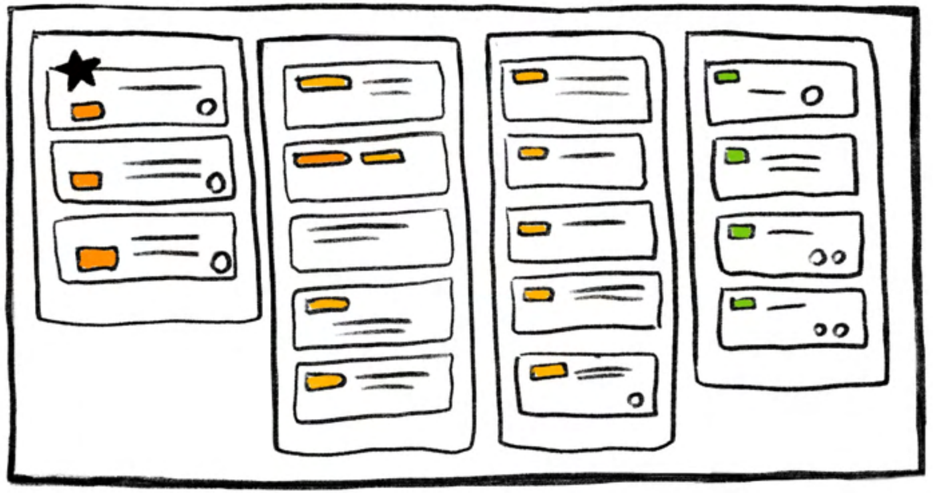
Shifting Left on Production Excellence with Observability

DevOps Enterprise Summit
October 5-7, 2021

featuring illustrations
by @emilywithcurls!

Your business wants to move faster



Velocity requires tight feedback loops

We need to shift production left.

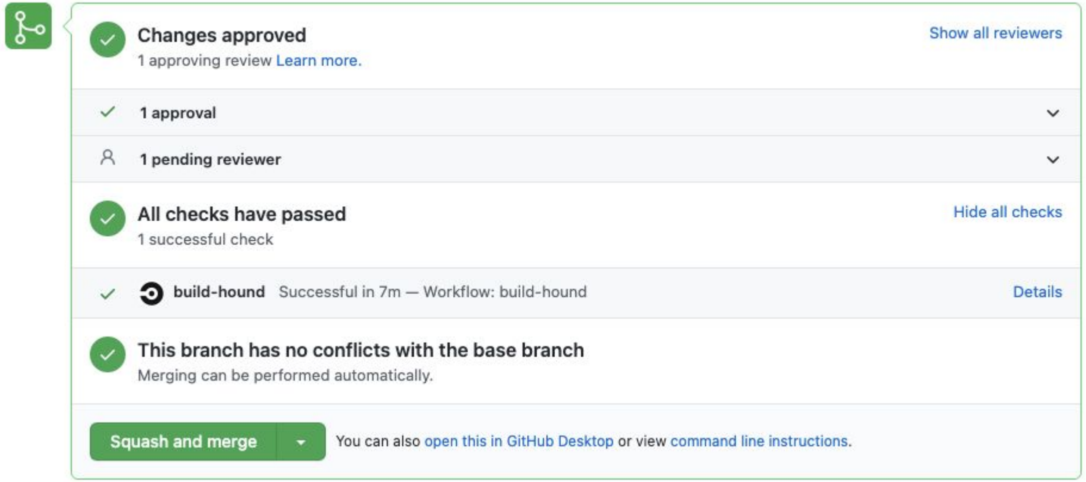
Liz Fong-Jones

Principal Developer Advocate at Honeycomb.io
lizthegrey.com
@lizthegrey

Shelby Spees

Site Reliability Engineer at Equinix Metal
spees.dev
@shelbyspees

We've come a long way



Changes approved
1 approving review Learn more.
Show all reviewers

✓ 1 approval

1 pending reviewer

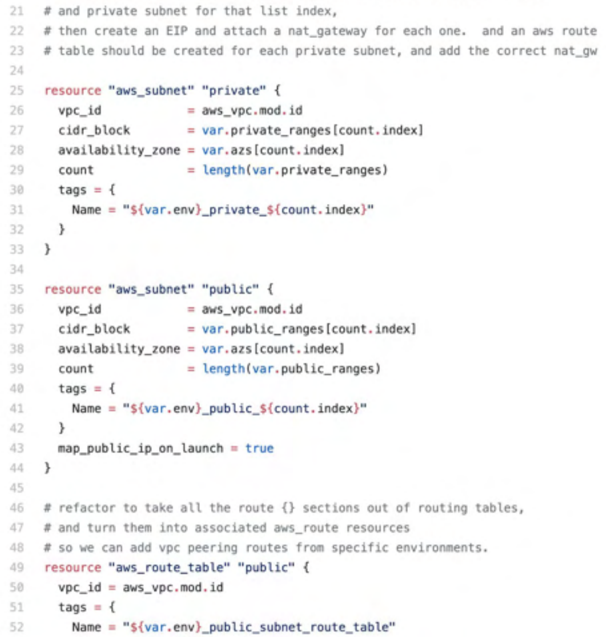
All checks have passed
1 successful check
Hide all checks

✓ build-hound Successful in 7m — Workflow: build-hound Details

This branch has no conflicts with the base branch
Merging can be performed automatically.

Squash and merge

You can also open this in GitHub Desktop or view command line instructions.



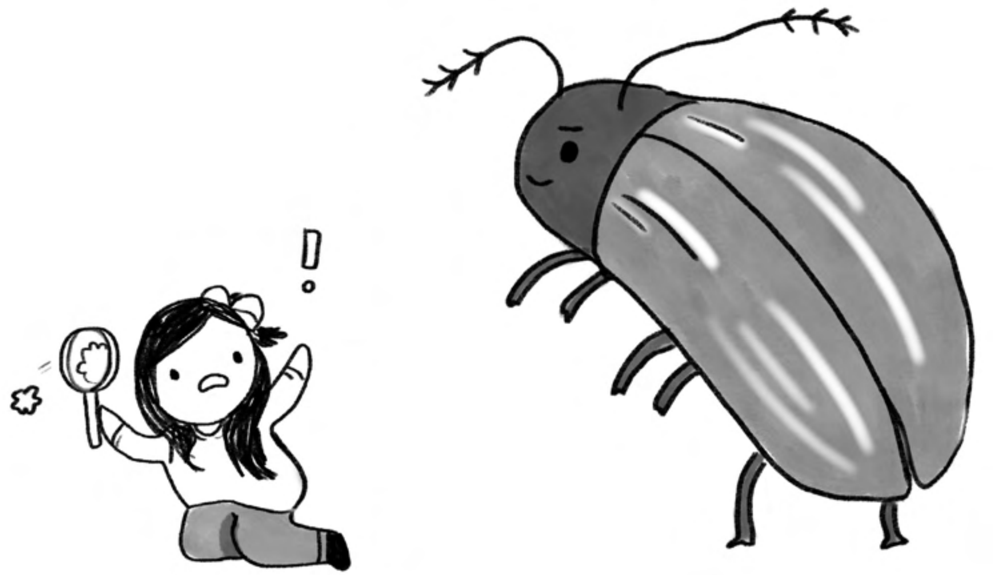
```
# and private subnet for that list index,
# then create an EIP and attach a nat_gateway for each one.  and an aws route
# table should be created for each private subnet, and add the correct nat_gw

resource "aws_subnet" "private" {
  vpc_id            = aws_vpc.mod.id
  cidr_block        = var.private_ranges[count.index]
  availability_zone = var.azs[count.index]
  count             = length(var.private_ranges)
  tags = {
    Name = "${var.env}_private_${count.index}"
  }
}

resource "aws_subnet" "public" {
  vpc_id            = aws_vpc.mod.id
  cidr_block        = var.public_ranges[count.index]
  availability_zone = var.azs[count.index]
  count             = length(var.public_ranges)
  tags = {
    Name = "${var.env}_public_${count.index}"
  }
  map_public_ip_on_launch = true
}

# refactor to take all the route {} sections out of routing tables,
# and turn them into associated aws_route resources
# so we can add vpc peering routes from specific environments.
resource "aws_route_table" "public" {
  vpc_id = aws_vpc.mod.id
  tags = {
    Name = "${var.env}_public_subnet_route_table"
```

What’s the next step?

Should developers be on-call?

Production is increasingly complex

Production feels intimidating

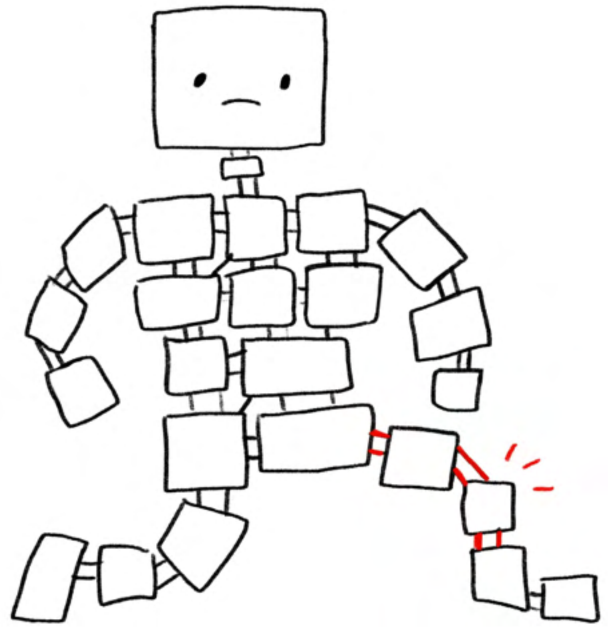


Traditional tools are inscrutable

Prod is always encountering new issues

emergent failure modes

small, unrelated failures cascading together to degrade or take down a system



see also: how.complexsystems.fail

42%

of developer time is spent dealing with bad code and tech debt

Source: [The Developer Coefficient, Stripe, 2018]

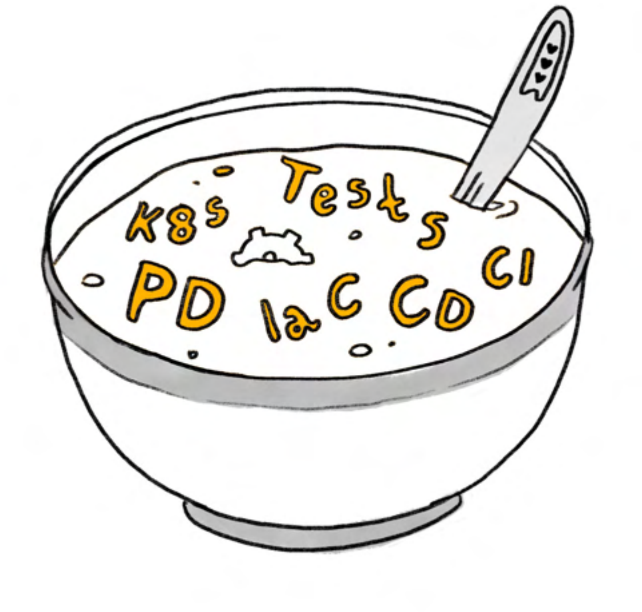
Teams can't make forward progress

Our heroes are exhausted

We shouldn’t stop here

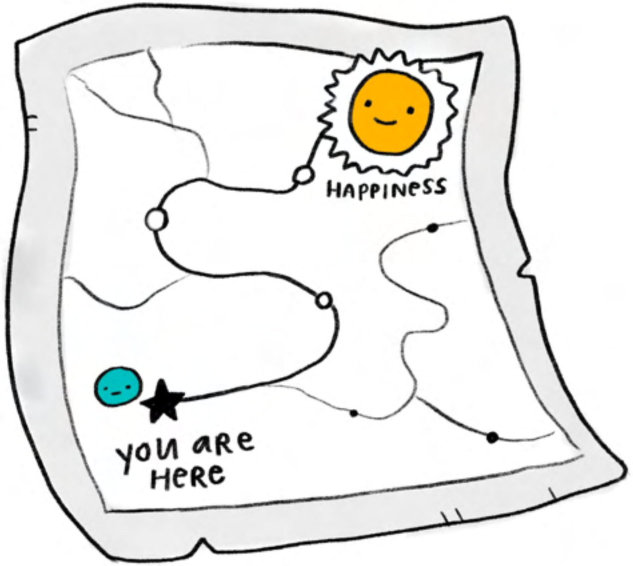
All teams need production excellence

Can't we just buy it?



Invest in people, culture, and process

Production Excellence is Business Excellence

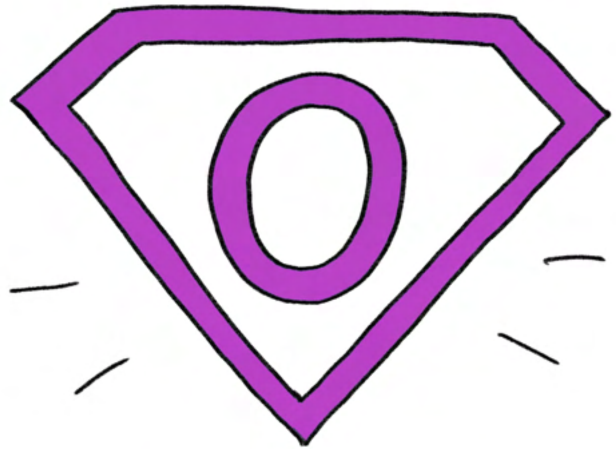


Start with observability

What is observability?

The ability to inspect and understand a system's internal state using the telemetry data it's already outputting.

(Even if your system is in a state you didn't know was possible!)



Hard-to-debug problems

Distributed Systems
small change causing downstream effects?

Poor Performance
what is worth optimizing?

Subset of Traffic
only some users are complaining?

Regulatory Requirements
can’t touch the box because of SOX?

Observability finds answers.

Distributed Systems
Not a problem, we have traces.

Poor Performance
Identify the bottleneck.

Subset of Traffic
Slice and dice with ad-hoc queries.

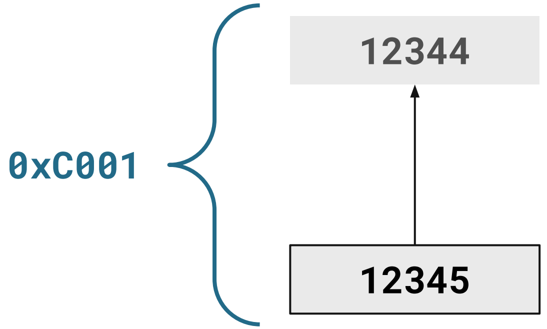
Regulatory Requirements
Read-only access.

Example event

```
event = {
    "duration_ms" = 1198,
    "request.id" = 12345,
    "request.method" = "GET",
    "request.path" = "/search",
    "request.query_string" = "category=decor&price<=50",
    "response.status_code" = 200
}
```

Example event

```
event = {
    "trace_id" = 0xC001,
    "span_id" = 12345,
    "parent_id" = 12344,
    "duration_ms" = 1198,
    "request.method" = "GET",
    "request.path" = "/search",
    "request.query_string" = "category=decor&price<=50",
    "response.status_code" = 200
}
```



Examples of dimensions to capture

infrastructure:

- build ID
- kubernetes pod
- kafka broker

application:

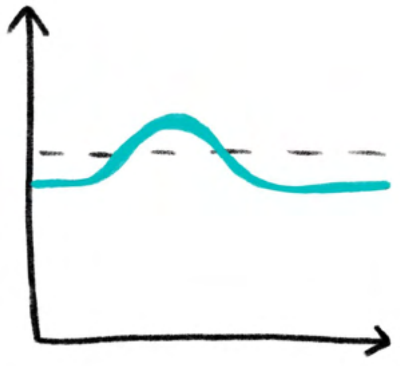
- library version
- user-agent
- sanitized SQL query string

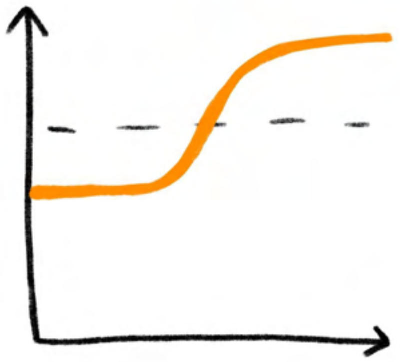
domain:

- number of items in cart
- coupon code
- payment processor (e.g. PayPal vs. Stripe)

Technical Decisions are Business Decisions

How do we map effort to impact?

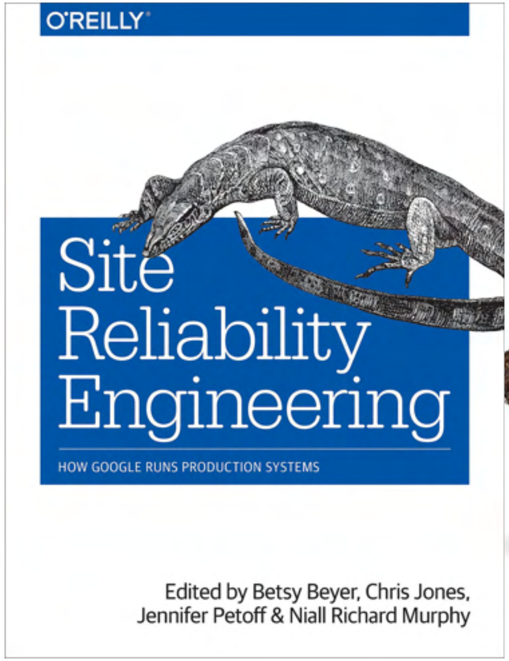




“How does this drive the business forward?”

Service Level Objectives (SLOs)

Common language for engineers and business stakeholders

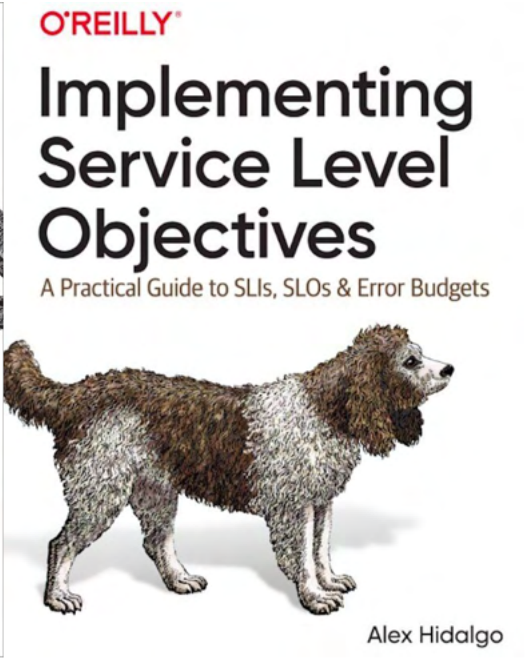


O'REILLY

Site Reliability Engineering

HOW GOOGLE RUNS PRODUCTION SYSTEMS

Edited by Betsy Beyer, Chris Jones, Jennifer Petoff & Niall Richard Murphy



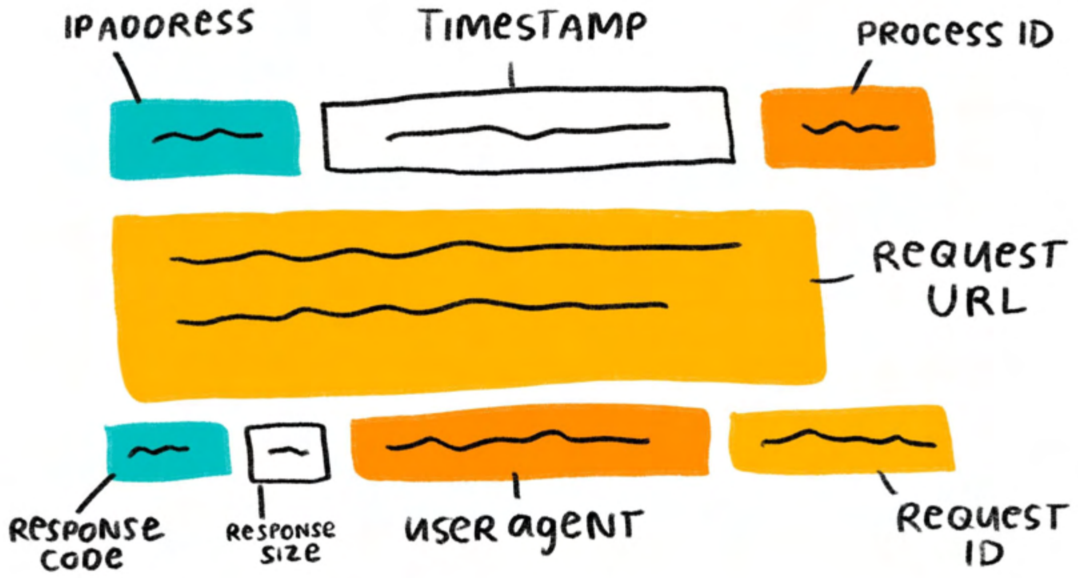
O'REILLY

Implementing Service Level Objectives

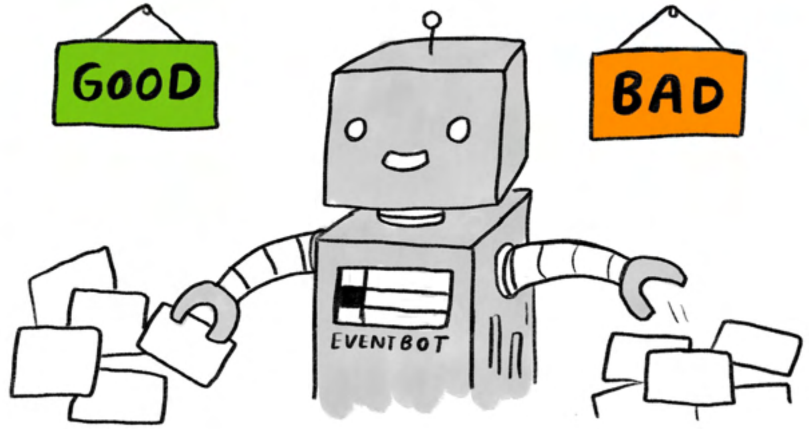
A Practical Guide to SLIs, SLOs & Error Budgets

Alex Hidalgo

Think in terms of events in context



SLI: for each event, good or bad?



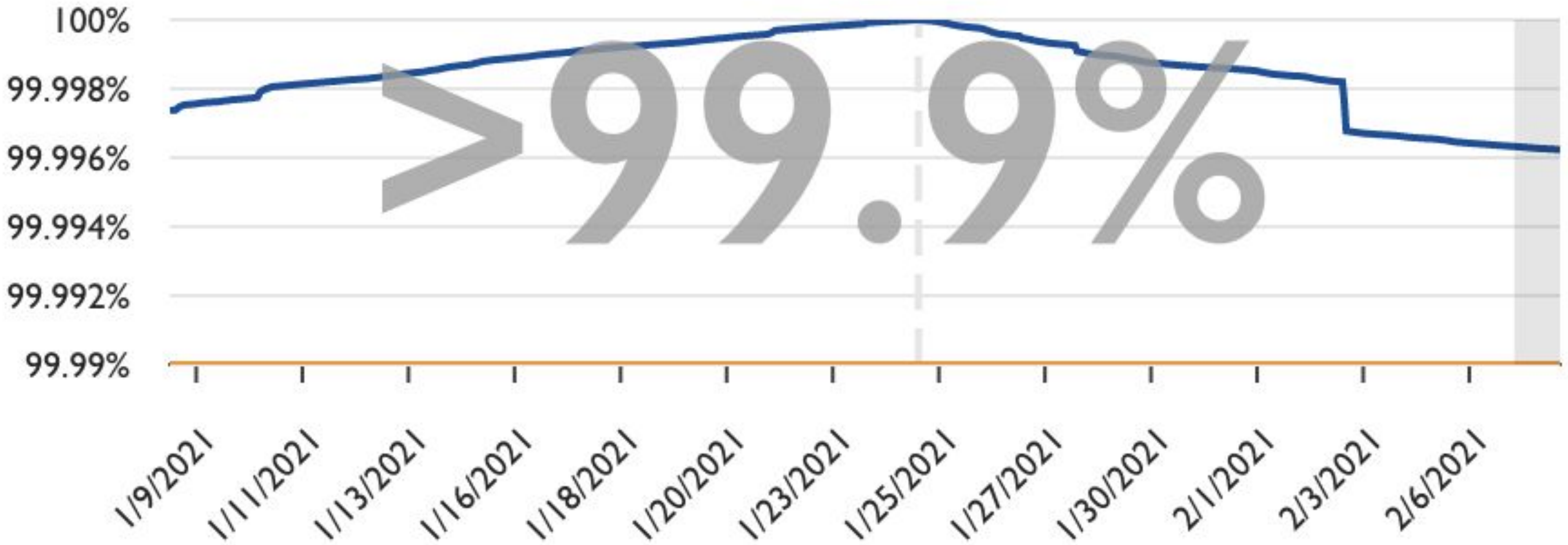
Use a window and target percentage

window is usually 30, 60, or 90 days

99.9%

of app Home Page loads
over the past 30 days
were “successful” and “fast enough”

Historical SLO compliance



A good SLO barely keeps users happy

Error budget: allowed unavailability

$$1 - x$$

99.9% SLO Target → 0.1% Error Budget

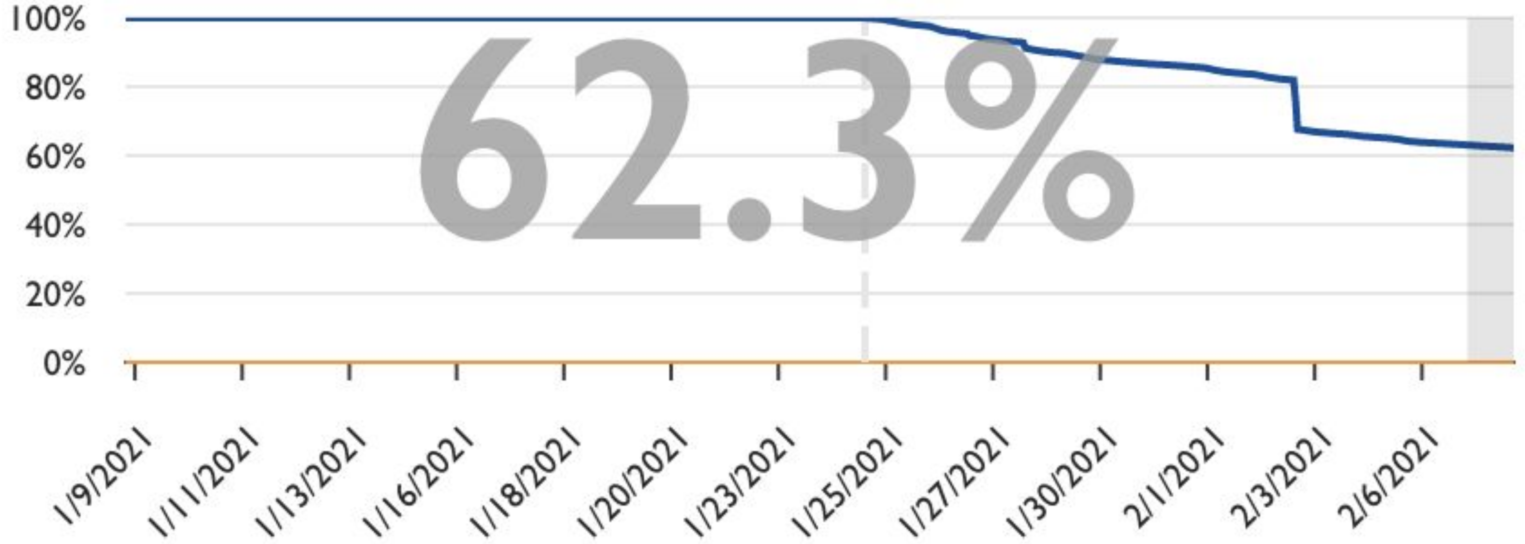
Example error budget

99.9% SLO Target → 0.1% Error Budget

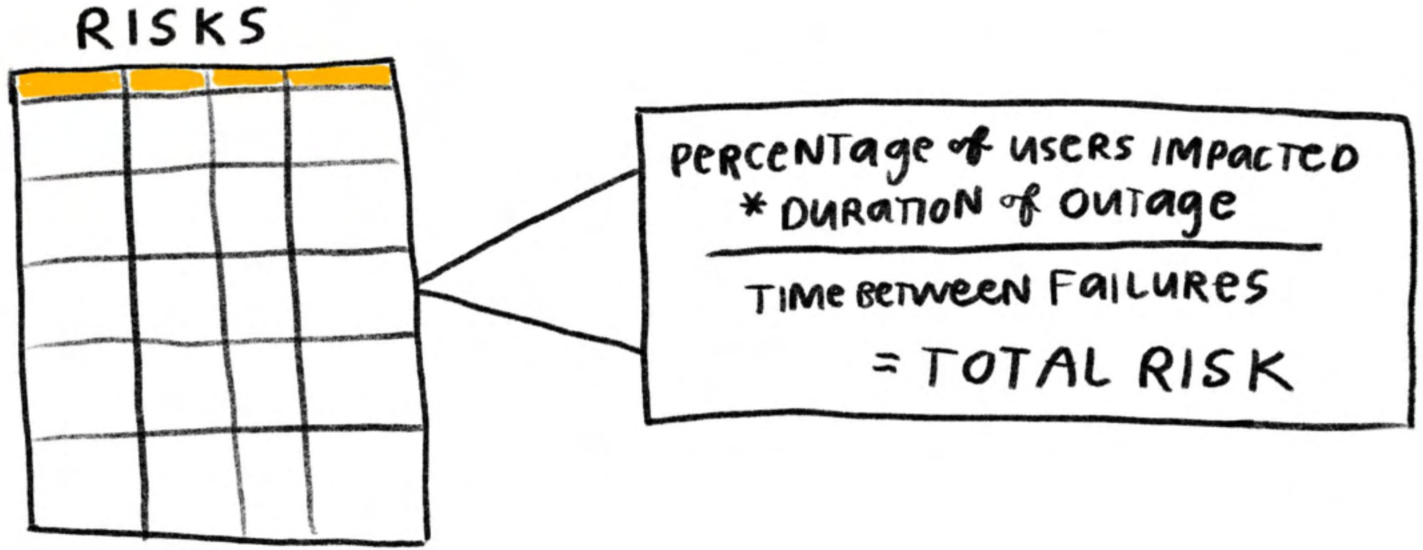
```
   1,000,000     requests/month
×        0.1%
──────────────
       1,000     requests/month
```

We’re allowed 1000 “bad” requests/month

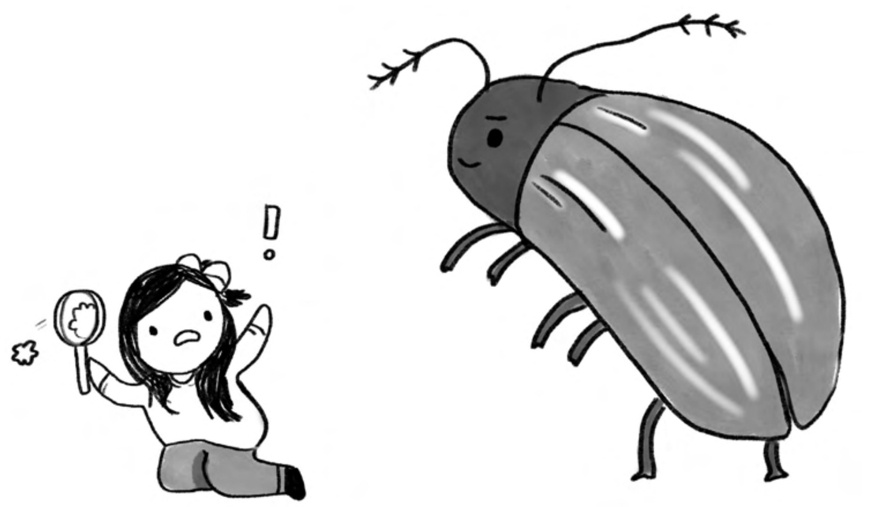
Error budget remaining

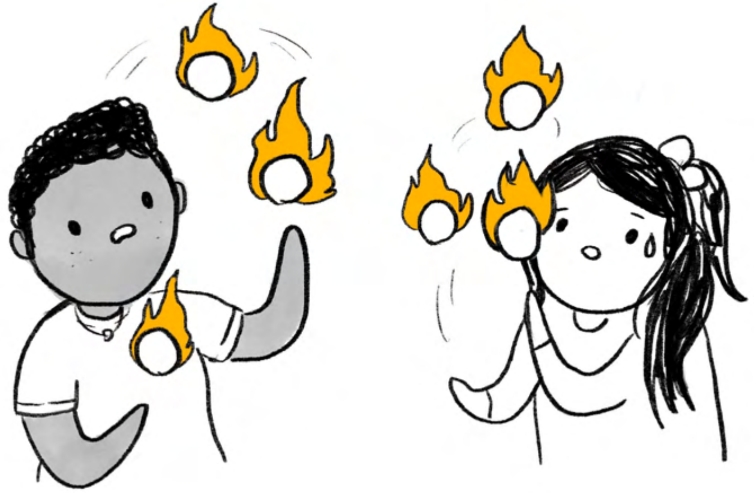


When is it okay to take risks?



When is it *not* okay?





Alerting on SLOs

Alert on error budget **burndown rate**

> “How soon are we going to run out of budget?”

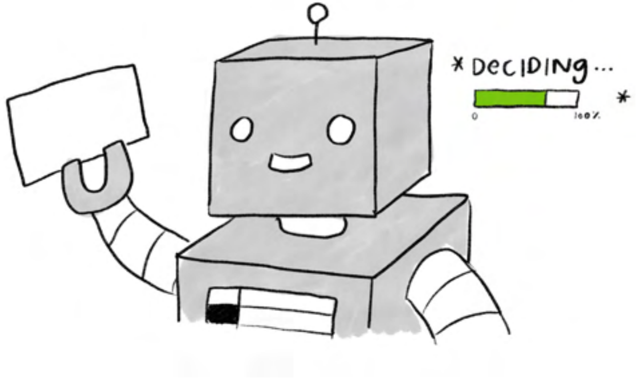
NOT: every **potential cause** of an alert

> “Disk is at 90.05%!!!” (does it matter?)

Observability & Continuous Delivery

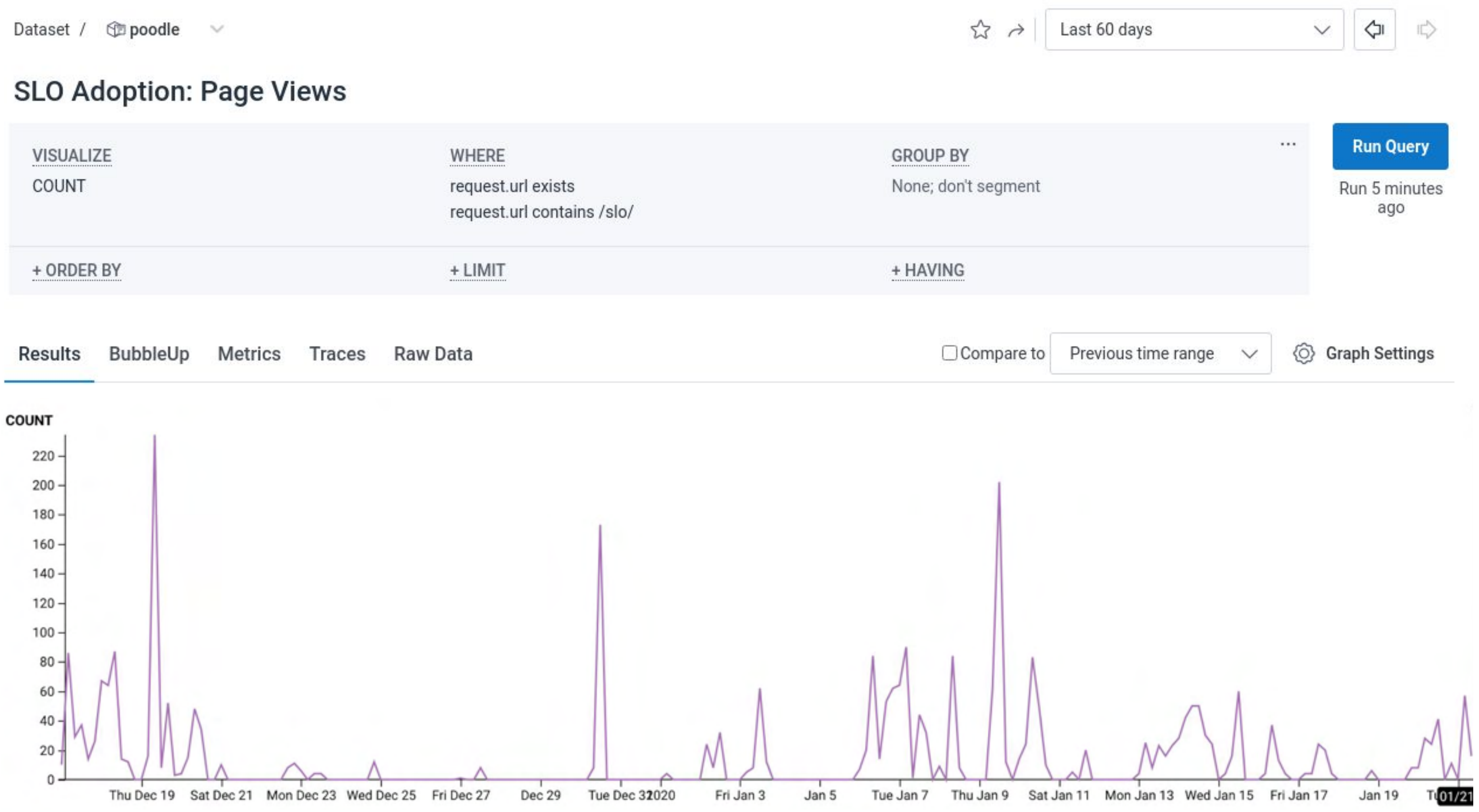
Continuous Delivery (at Honeycomb)

Instrument as we code.



Ensure fast integration.

Observe behavior in prod.



Dataset / poodle

Last 60 days

SLO Adoption: Page Views

VISUALIZE	WHERE	GROUP BY
COUNT	request.url exists request.url contains /slo/	None; don't segment

+ ORDER BY + LIMIT + HAVING

Run Query

Run 5 minutes ago

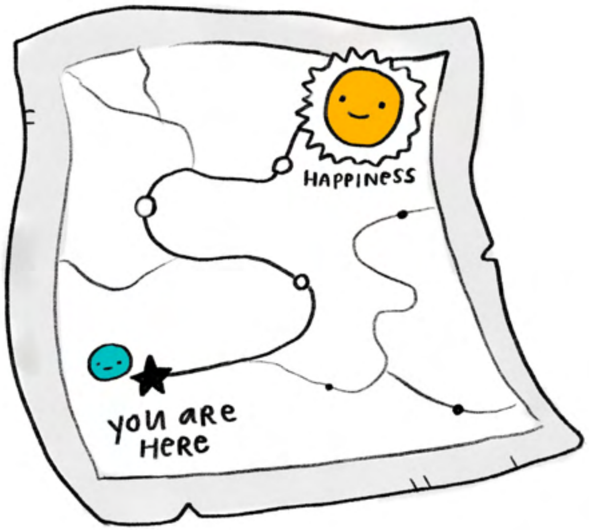
Results BubbleUp Metrics Traces Raw Data

Compare to Previous time range

Graph Settings

Bring production excellence to your org

Invest in the right places



Lean on auto-instrumentation

OpenTelemetry

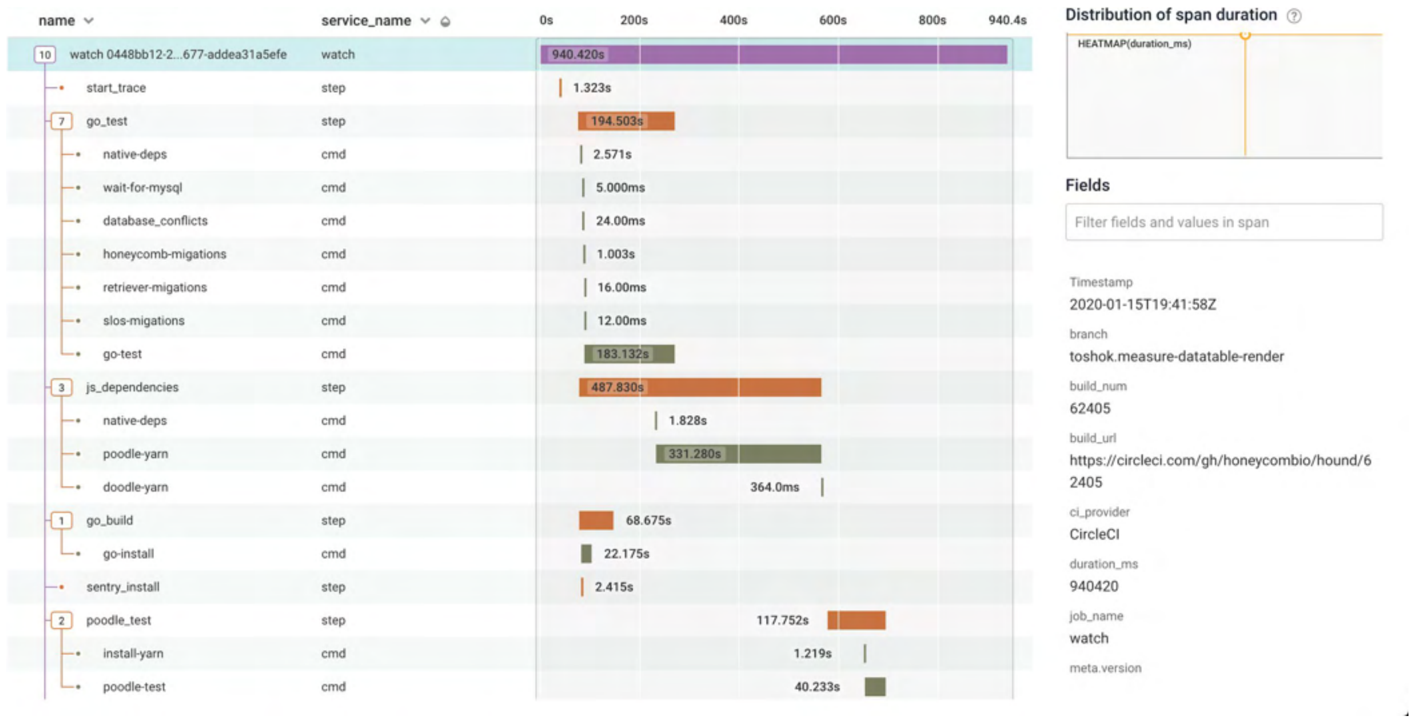
Java

Microsoft .NET

node JS

python

Instrument your builds



name	service_name	0s 200s 400s 600s 800s 940.4s
10 watch 0448bb12-2...677-addea31a5efe	watch	940.420s
start_trace	step	1.323s
7 go_test	step	194.503s
native-deps	cmd	2.571s
wait-for-mysql	cmd	5.000ms
database_conflicts	cmd	24.00ms
honeycomb-migations	cmd	1.003s
retriever-migations	cmd	16.00ms
slos-migations	cmd	12.00ms
go-test	cmd	183.132s
3 js_dependencies	step	487.830s
native-deps	cmd	1.828s
poodle-yarn	cmd	331.280s
doodle-yarn	cmd	364.0ms
1 go_build	step	68.675s
go-install	cmd	22.175s
sentry_install	step	2.415s
2 poodle_test	step	117.752s
install-yarn	cmd	1.219s
poodle-test	cmd	40.233s

Distribution of span duration

Fields

Filter fields and values in span

Timestamp	2020-01-15T19:41:58Z
branch	toshok.measure-datatable-render
build_num	62405
build_url	https://circleci.com/gh/honeycombio/hound/62405
ci_provider	CircleCI
duration_ms	940420
job_name	watch
meta.version	

Find a champion

Learn and iterate

SLOs aren't one-and-done

- Adjusting SLO targets
- Improving instrumentation
- Updating alerting methods

see also: [go.hny.co/biz-goals-slos]

Make it scale to your org

Prioritize:

- developer experience
- on-call onboarding

Invest in:

- config as code
- custom instrumentation libraries

Case Study: Observability at Vanguard

Observability champion(s) ✔️

Prioritizing knowledge transfer ✔️

OpenTelemetry ✔️

Service Level Objectives ✔️

18 million developers

It’s on us to grow them into production engineers.

Help we're looking for

CNCF Observability
Technical Advisory Group (TAG)

- CNCF Slack **`#tag-observability`** channel

OpenTelemetry

- CNCF Slack **`#opentelemetry`** channel
- OpenTelemetry.io

OpenSLO

- OpenSLO Slack
- OpenSLO.com

Schedule 30 minutes

- hny.co/meet/liz

Find us on Twitter

- twitter.com/lizthegrey
- twitter.com/shelbyspees