Thawing the Great Code Slush

SRECon Americas 2024

Maude Lemaire / @qcmaude

Sr. Staff Engineer @ Slack
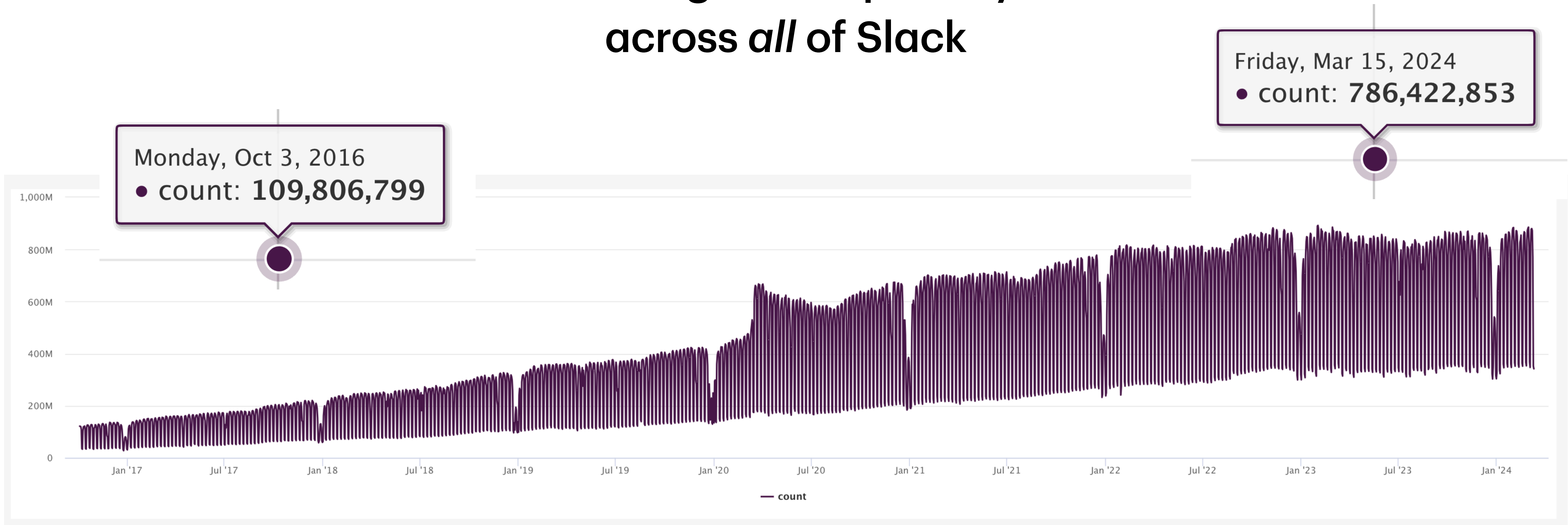
Hometown is the birthplace of poutine
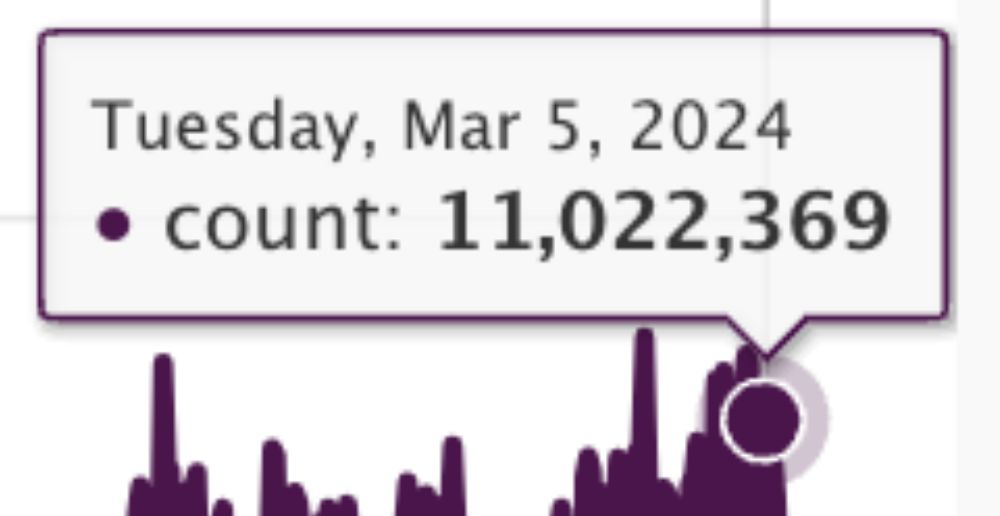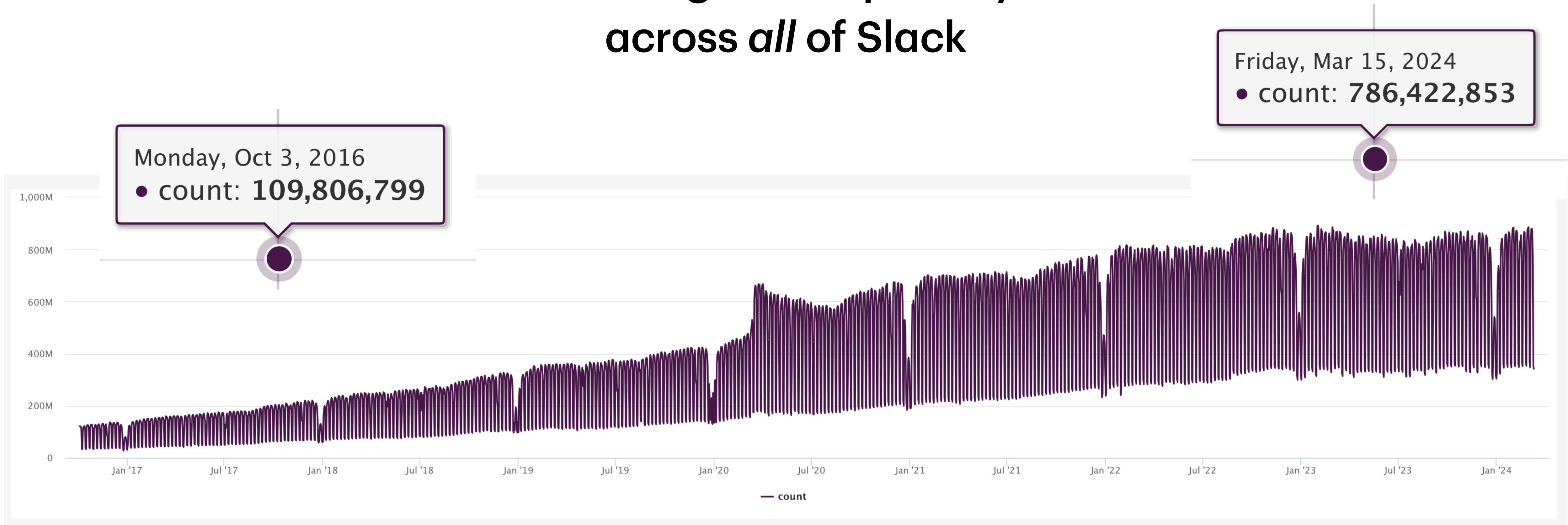
* Drummondville, QC

~~Victoriaville~~

Messages sent per day across *all* of Slack

Monday, Oct 3, 2016
● count: **109,806,799**

Friday, Mar 15, 2024
● count: **786,422,853**

Messages sent per day across *all* of Slack

Monday, Oct 3, 2016
● count: **109,806,799**

Friday, Mar 15, 2024
● count: **786,422,853**

Tuesday, Mar 5, 2024
● count: **11,022,369**

I've worked on:

BIG refactors

Enterprise Grid™

👉 **(backend) performance**

HHVM

core libraries

load testing 👉

October 2020

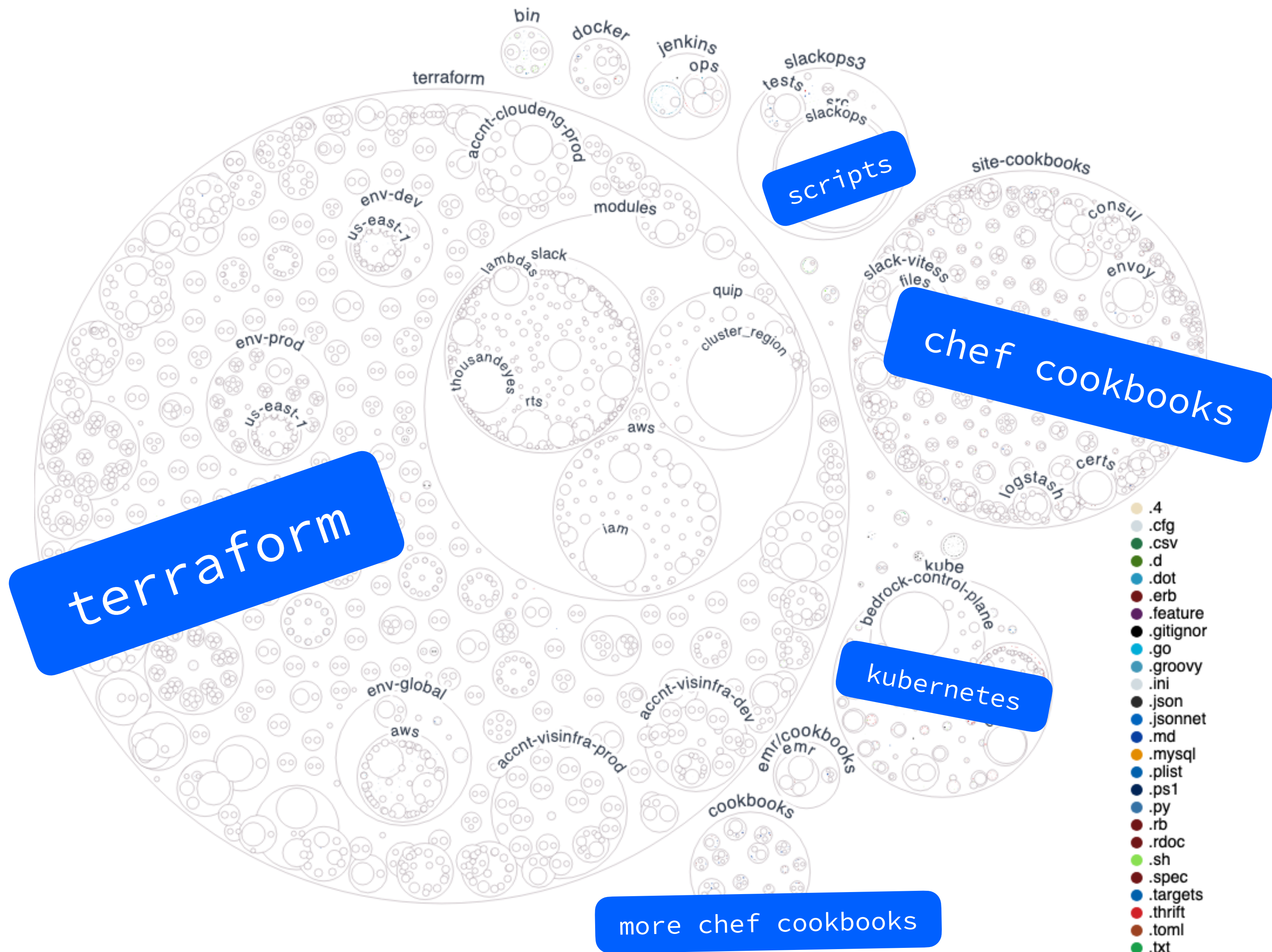Our quarterly uptime, assuming 100% for the rest of October, will be 99.96%.
That is the lowest quarterly uptime in over a year.

- On October 2nd, a new HHVM build was promoted. A subsequent deployment of the Slack webapp (run atop HHVM), caused a 4h SEV-1 incident, where we saw an elevated spike in 500s to our web API.

- On October 5th, we had a significant decrease in API success rate that lasted over a day and a half (39 hours). First, we discovered problems with our memcache cluster, and then consul struggling under heavy load. We eventually figured out that the service discovery issues were related to a percentage-based roll-out of our service discovery cache.

- On October 16th, Slack was degraded for all users for ~13 hours. Memcache stopped working on many hosts, causing crushing loads on our databases. The root cause was a change to mcrouter 8 days prior.

- On October 2nd, a new HHVM build was promoted. A sub[...]nent of the Slack webapp (run atop HHVM), caused a 4h SEV-1 in[...]evated spike in 500s to our web API.

- On October 5th, we had a[...]ss rate that lasted over a day and a half (39 h[...] with our memcache cluster, and then consul strugg[...]entually figured out that the service discovery issues were rel[...]ge-based roll-out of our service discovery cache.

- On October 16th[...]ck was degraded for all users for ~13 hours. Memcache stopped working on many hosts, causing crushing loads on our databases. The root cause was a change to mcrouter 8 days prior.

chef-repo

bin
docker
jenkins
ops
terraform
slackops3
tests
src
slackops
scripts
accnt-cloudeng-prod
env-dev
site-cookbooks
us-east-1
modules
consul
slack
slack-vitess
envoy
lambdas
files
quip
thousandeyes
cluster_region
env-prod
rts
us-east-1
aws
certs
logstash
iam
chef cookbooks
terraform
kube
bedrock-control-plane
env-global
kubernetes
aws
accnt-visinfra-dev
emr/cookbooks
accnt-visinfra-prod
emr
cookbooks
more chef cookbooks

.4
.cfg
.csv
.d
.dot
.erb
.feature
.gitignor
.go
.groovy
.ini
.json
.jsonnet
.md
.mysql
.plist
.ps1
.py
.rb
.rdoc
.sh
.spec
.targets
.thrift
.toml
.txt
.webapp
.xml
.yaml
.yml
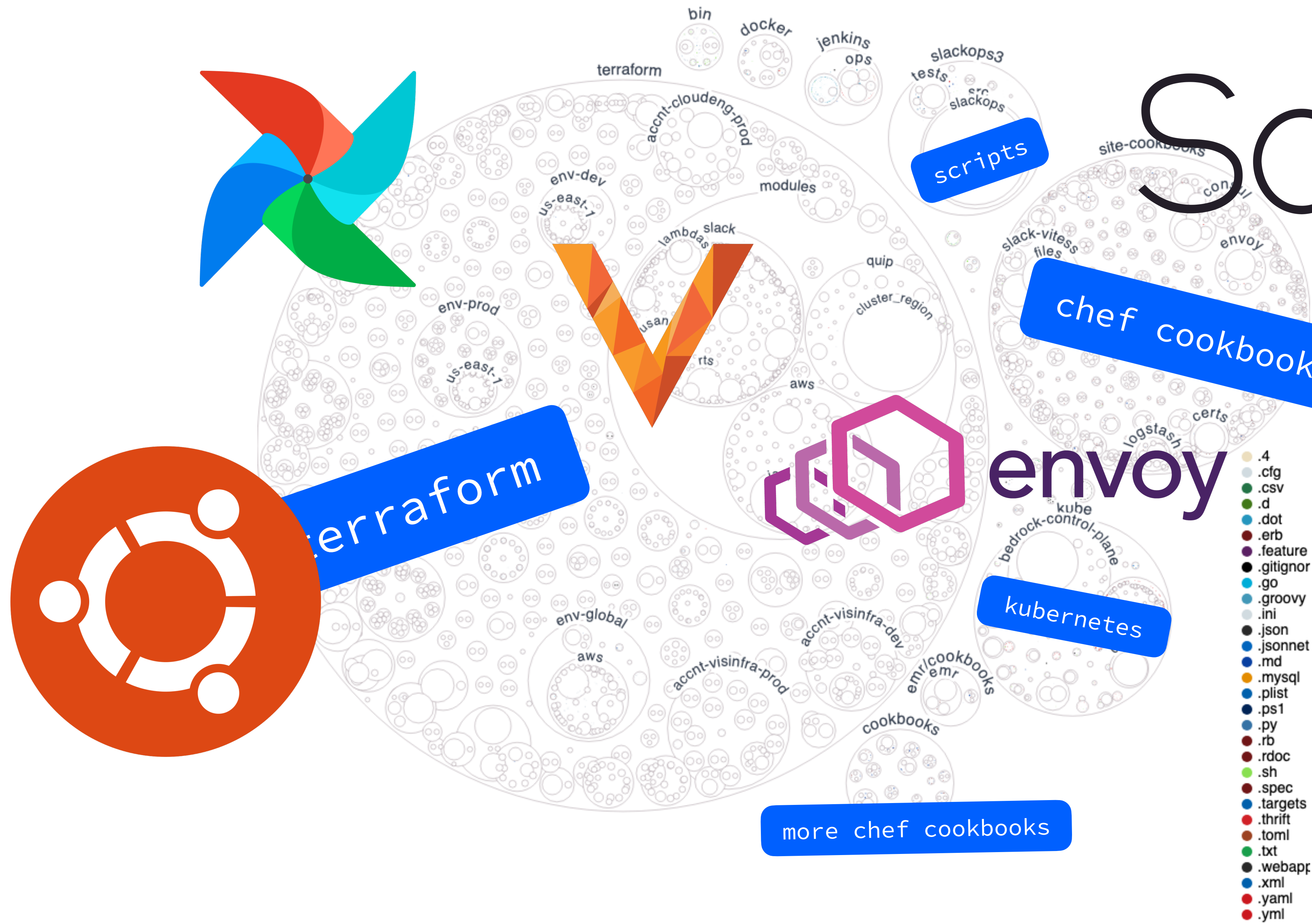
Nope 🛑 ✋

"We have established a **Change Advisory Board (CAB)** that meets **nine** times a week (twice daily in the mornings and afternoons, except Friday afternoon) to review proposed changes.

Changes are submitted in #infra-change-control and we require the requestor to manually monitor any approved changes and provide status and visibility updates in channel."
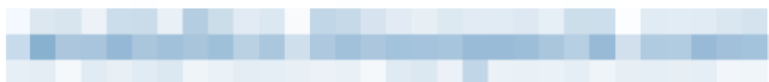
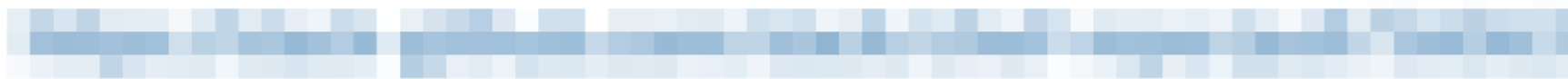**Change Request** `FLUX DE TRAVAIL` 10 h 31

Merge Approval Request submission from ▓▓

What service/feature being changed?
Presto

What does your change do?
Adds additional ▓▓▓▓ ▓▓▓▓▓▓ ▓▓ dashboards into a high-concurrency worker pool on Presto.

URL to your code-reviewed PR
▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓

Why is this a high-priority change?
Support Q3 goals towards improving VIP analytics dashboards - a proven 10x improvement.

What is the risk level of this change?
Medium

Checklist questions answered in PR description?
Yes

▓▓ **6 réponses** Dernière réponse il y a 3 ans

Who were these people?

Change Request FLUX DE TRAVAIL 10 h 31
Merge Approval Request submission from ▓▓

What service/feature being changed?
Presto

What does your change do?
Adds additional ▓▓▓▓ ▓▓▓▓ ▓▓ dashboards into a high-concurrency worker pool on Presto.

URL to your code-reviewed PR
▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓▓

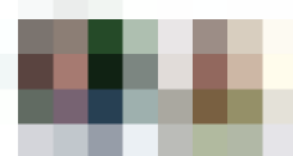**Why is this a high-priority change?**
Support Q3 goals towards improving VIP analytics dashboards - a proven 10x improvement.

What is the risk level of this change?
Medium

Checklist questions answered in PR description?
Yes

▓▓▓ **6 réponses** Dernière réponse il y a 3 ans

#infrastructure–safety

- **Reduce risk** associated with memcache, consul, other system-critical infrastructure

- Add automated testing via test kitchen

- Enable gradual rollouts for a wider breadth of changes

- **Reduce risk** associated with memcache, consul, and other system-critical infrastructure

LOW MEDIUM HIGH

- **Reduce risk** associated with memcache, consul, and other system-critical infrastructure

OBSERVABILITY   RESILIENCY   READINESS

ABILITY TO RESCUE   BLAST RADIUS OF DEPLOY

- **Add automated testing via test kitchen**

  * <u>Brett & Leigh's Talk</u>

- **Enable gradual rollouts for a wider breadth of changes**

- **Reduce risk** associated with memcache, consul, and other system-critical infrastructure

- **Add automated testing via test kitchen**

- **Enable gradual rollouts for a wider breadth of changes**

- **Reduce risk** associated with memcache other system criti... for a wider breadth of changes

How long is the freeze window
- We do not know how long it will take to create a new safe permanent process, but 4 weeks is a reasonable estimate.

7 months

	chef-repo	
Distinct contributor count (June 2020– June 2021)	**335**	
Distinct contributor count (Month preceding slush)	**170**	
Total PR count (Month preceding slush)	**2,473**	

	chef-repo	webapp
Distinct contributor count (June 2020– June 2021)	335	744
Distinct contributor count (Month preceding slush)	170	407
Total PR count (Month preceding slush)	2,473	6,944

¯_(ツ)_/¯

March 16, 2021

Management 🤠

- **Accountability**
- **Understand the risk profile of changes being merged**

Everyone ✨

- **Trust**
- **Reliability**
- **Time** ⏰

Engineers 😎

- **Velocity**
- **Independence**

Management 🤠   False Dichotomy   **Engineers** 😎

Management 🤠

We can only* release safe, reliable code if it's properly vetted and approved by domain experts and engineering leadership.

*At the very least, until we have the necessary safety mechanisms in place, including better monitoring, unit testing, greater ability to deploy incrementally, etc.

Engineers 😎

Engineering velocity should be prioritized to ensure safety and reliability.

* If we slow down development, and most importantly, deployments, then our muscles around quick iteration will atrophy, and we'll ultimately end up with a less reliable product because we won't be able to fix problems quickly.

Isn't CAB essentially just ... *really good* code review?

What makes good code review?

`#24amer-day2-track1`

Why don't *you* build it?

Filters ▾   🔍 author:maude is:closed merged:<2021-03-16   🏷 **Labels** `222`   ⚑ **Milestones** `0`   **New issue**

✕ **Clear current search query, filters, and sorts**

| ☐ ⊙ 0 Open ✓ **9 Closed** | Author ▾   Label ▾   Projects ▾   Milestones ▾   Assignee ▾   Sort ▾ |

☐   **Slerf is now a Bedrock service!** ●   💬 1
     #60939 by maude was merged on Mar 1, 2021 • Review required

☐   **Add `koi_pond.yaml` to production loadtest cluster!** ●
     #53712 by maude was merged on Sep 4, 2020 • Review required

☐   **Add `puppet-show` to list of Hound repos** ●
     #49934 by maude was merged on Jun 23, 2020 • Approved

☐   **Remove Derek from #escal-backend-foundation** ●   💬 1
     #44246 by maude was merged on Mar 3, 2020 • Approved

☐   **Roll out to production (except polyglot) for 4.27!** ●   💬 2
     #41173 by maude was merged on Dec 18, 2019 • Approved

☐   **Roll out HHVM 4.27 to dogfood and canary!** ●   💬 3
     #41116 by maude was merged on Dec 17, 2019 • Approved

☐   **Roll out HHVM 4.27 to dev mains and qa multis!** ●   💬 3
     #41097 by maude was merged on Dec 17, 2019 • Review required

☐   **Moving Slerf from BBS to K8S** ●
     #39164 by maude was merged on Dec 5, 2019 • Approved

☐   **Specify grants on reading the `channels` table** ●   💬 1
     #14064 by maude was merged on Dec 13, 2017 • Review required

9 PRs

Beginner's Mind 🧠

Social Capital 💰

CAB Deputies

 CAB Deputies

- **Build code review training accessible to all committers of** `chef-repo`**.**

- **Create an assessment to test engineers' CAB review skills.**

- **Teach and onboard progressively larger groups of engineers.**

- **Distribute the load! At least 2** deputies per team (product area) **distributed regionally** (optimal turnaround for PR review).

- **Reinforce** engineers' understanding of `chef-repo` **shenanigans.**

- Bring a **renewed curiosity** to code review.

Everyone ✨
- **Trust**
- **Reliability**
- **Time** ⏰

Engineers 😎
- **Velocity**
- **Independence**

- **Free up** engineering management time!

- Deliver **accurate aggregate data** on the kinds of PRs merged, and **timely insights into the kinds of incidents** rooted in a `chef-repo` **PR.**

- Bring a **renewed accountability** to code review.

Everyone ✨

- **Trust**
- **Reliability**
- **Time** ⏰

Management 🤠

- **Accountability**
- **Understand the risk profile of changes being merged**

talk for roughly 35 min... (You can also tweak the replay ...

AuZYVn9*. (You can also tweak the replay ...

time.)

Training

Current CAB Goal ...

CAB seeks to address the fo... ...ikely through a

- **Blast Radius:** Can the blast radius of this ...percentage based roll out.
- **Testability:** With what degree of fidelity was this change tested?
- **Observability:** What metrics and dashboards will you be watching when this rolls out?
- **Rescue Plan:** Is there a rollback PR and plan? Can we feature flag (FF) the change?

Deputies Goals

At the end of this training, we want new CAB deputies to have a solid understanding of what makes a change low or standard risk – basically, you all become CAB approvers, but distributed across your teams. As part of this training, you'll solidify your understanding of *how* changes are rolled out via chef and *what* the blast radius of a change might look like throughout its deployment.

As Brett showed us in the Core Infra All Hands (Q1 2021), our chef dependencies are complicated.

"Probably, unless [you're] superhuman, [you're] not going to know everywhere that uses [the `chef-repo` code you're modifying] and test it everywhere."

Training

CAB seeks to address the following concerns and questions:

- **Blast Radius:** Can the blast radius of this deploy be reduced? Likely through a percentage based roll out.

- **Testability:** With what degree of fidelity was this change tested?

- **Observability:** What metrics and dashboards will you be watching when this rolls out?

- **Rescue Plan:** Is there a rollback PR and plan? Can we feature flag (FF) the change?

Training

CAB seeks to address the following concerns and questions:

- **Blast Radius:** Can the blast radius of this deploy be reduced?

OBSERVABILITY   RESILIENCY   READINESS

- **Testability:** With what degree of fidelity was this change tested?

BLAST RADIUS OF DEPLOY   ABIILITY TO RESCUE
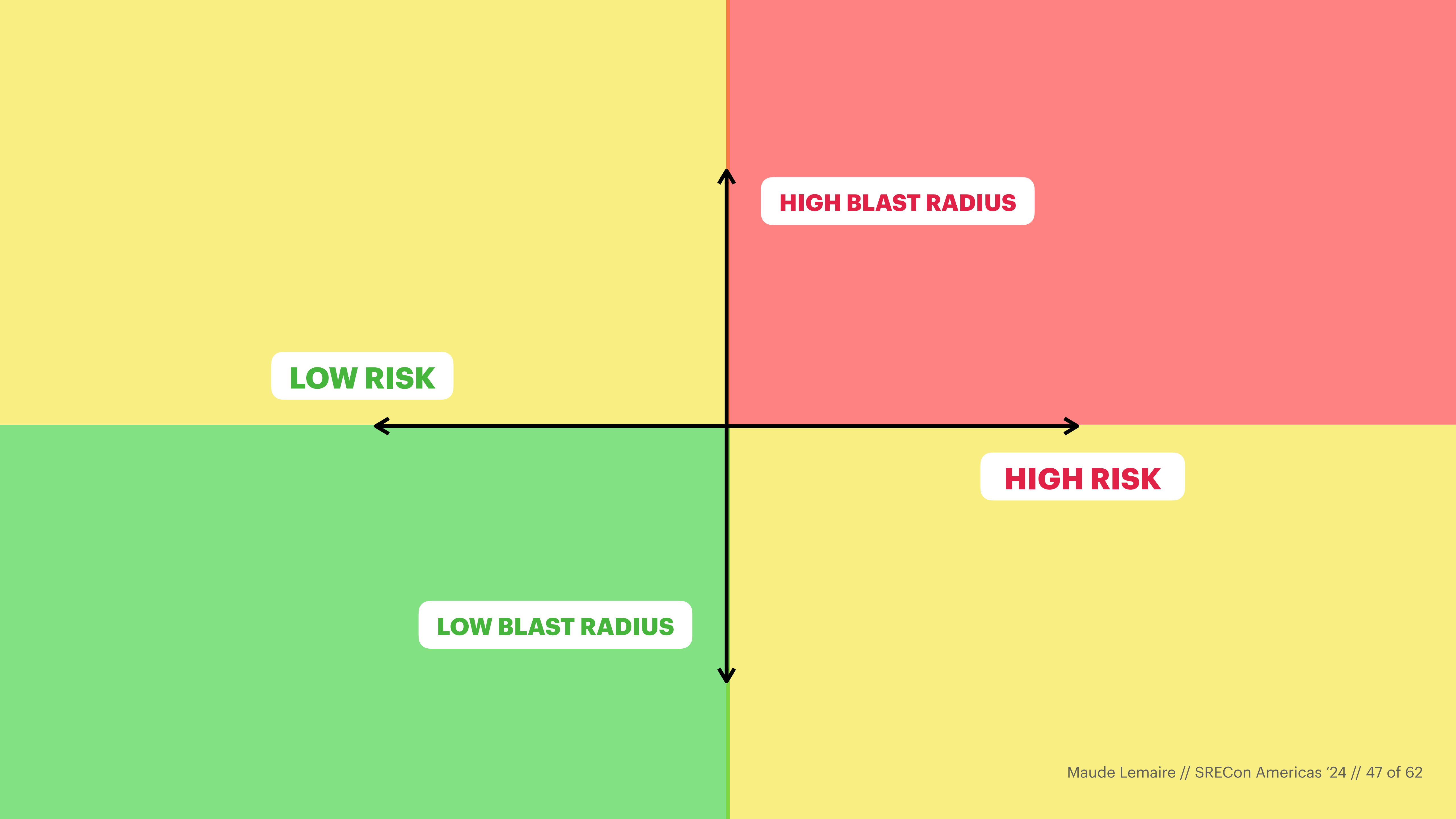
watching when this rolls out?

- **Rescue Plan:** Is there a rollback PR and plan? Can we feature flag (FF) the change?

🛡️ Training

- **Chef & chef-repo architecture refresh**

- **How to determine risk level and blast radius**

- **Logistics (i.e. how to properly tag and approve a PR)**

- **Example PRs (with corresponding risk and blast radius profiles)**

- **Example questions to ask code author in the PR**

- **Quiz!**
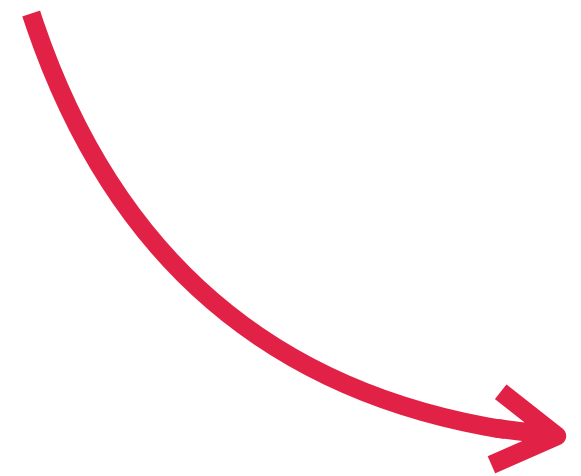
HIGH BLAST RADIUS

LOW RISK

HIGH RISK

LOW BLAST RADIUS

⬡ Training

- **Chef & chef-repo architecture refresh**

- **How to determine risk level and blast radius**

- **Logistics (i.e. how to properly tag and approve a PR)**

- **Example PRs (with corresponding risk and blast radius profiles)**

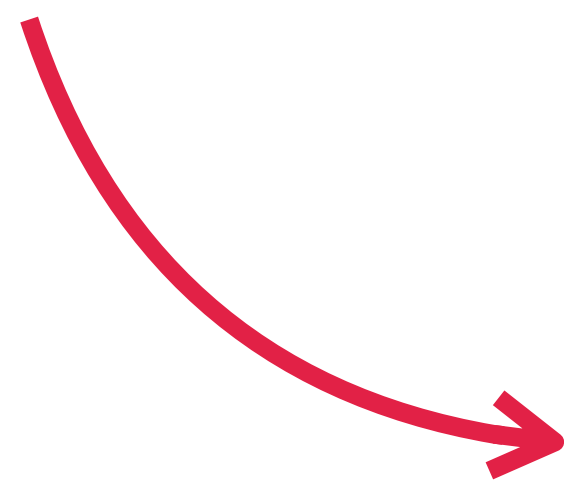- **Example questions to ask code author in the PR**

- **Quiz!**

What makes good code review?

What makes good code review?

* **Glen's take**

rosew.dev   3w ...

Beware of Staff+ Developers asking simple questions.

♡   ◯   ⇄   ◁

18 replies · 111 likes

- How is this change rolled out?
- What's the worst-case blast-radius associated with this change?
- What's the absolutely worst-case outcome that this change could have?
- How have you tested this change? What level of confidence has these test(s) given you about your change?
- How is it being monitored (both during rollout and after)?
- What dashboard (and panels, specifically) are you watching to make sure this change has the intended effect?
- If you're creating a new _____ for rollout and ultimately *removing* the flag?
- How will you know if this

Questions

- Does this change go to GovSlack?
    - Note that FF by default applies to both commercial and gov.
    - Make sure the rollout plan contains how to deploy changes to both Gov and Commercial.
    - Are any dependencies for the change present in both gov and commercial?
- How could you make this change even more safely (even if it was slower)?
- How often have you made a change of this nature? How routine is this change in general? How can you continue to improve its automation and safeguards?
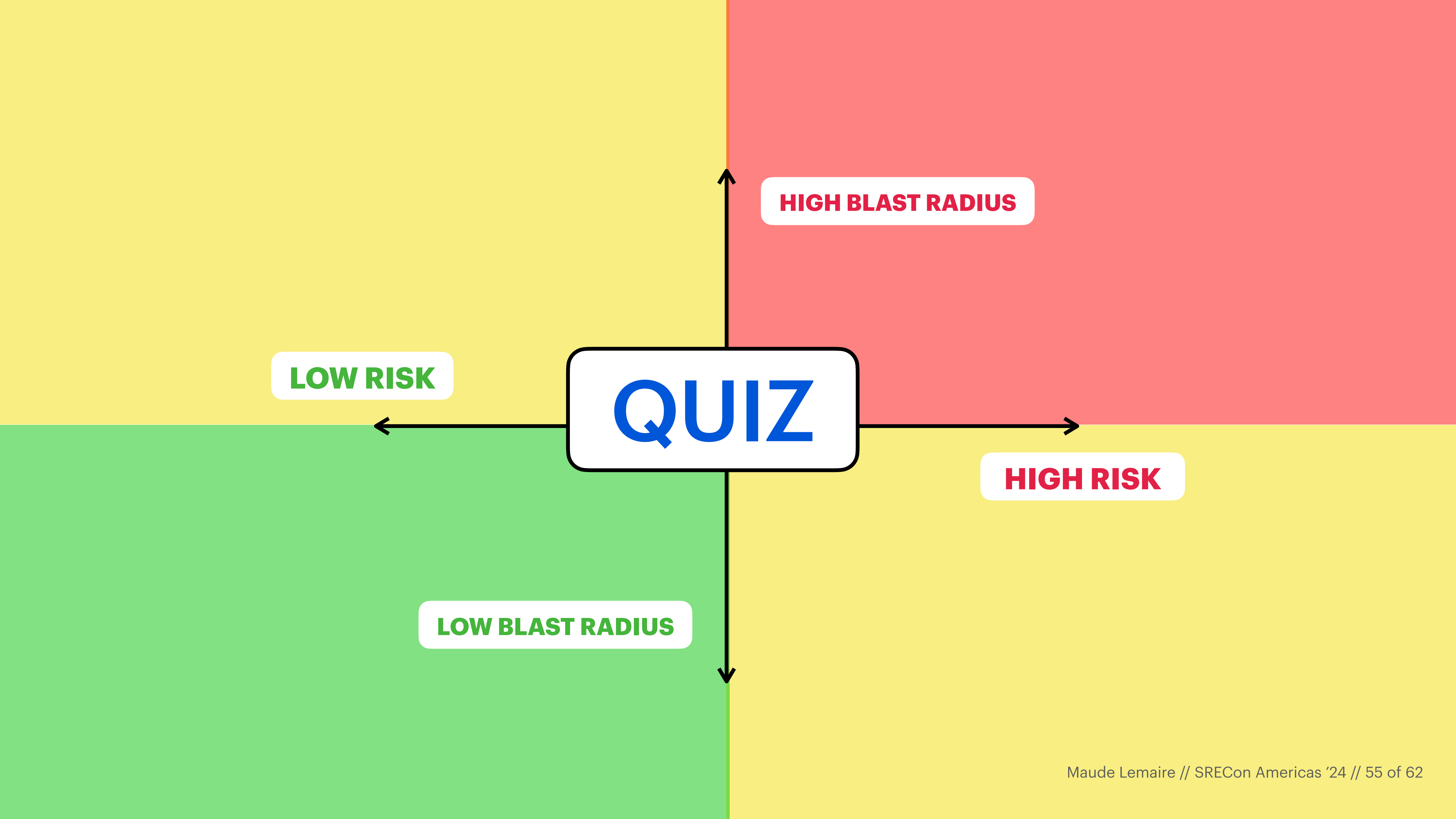
Training

- **Chef & chef-repo** architecture refresh

- **How to determine** risk level **and** blast radius

- **Logistics** (i.e. how to properly tag and approve a PR)

- **Example PRs** (with corresponding risk and blast radius profiles)

- **Example questions** to ask code author in the PR

- **Quiz**!

QUIZ

HIGH BLAST RADIUS

LOW RISK

QUIZ

HIGH RISK
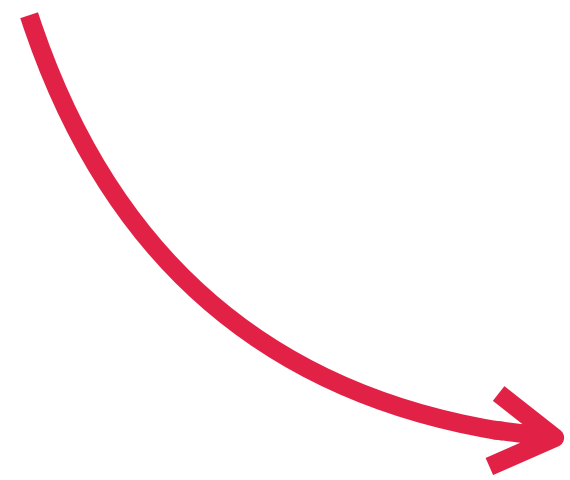
LOW BLAST RADIUS

Logistics

🏷️ 👤 **maude** added `cab-dep-low-blast-radius` `cab-dep-low-risk` labels on Aug 25, 2021

👁️ **maude** previously approved these changes on Aug 25, 2021   **View reviewed changes**

maude left a comment   (Member) ☺ • • •

This is a low risk, low blast-radius change that only impacts Koi Pond S3 buckets in our own account. TF plans look good.

Remind me which pipeline is this built under?

Logistics

✓ **gsanford** approved these changes on behalf of **slack/datastore** on May 25, 2021   **View reviewed changes**

gsanford left a comment   Member 😊 ⋯

Thanks for cleaning this up.

CAB:

- standard risk (changes to chef code to remove FF)
- high blast radius: *technically* this affects 1/3 of the running webapps, but it affects a variant of mcrouter that isn't being used right now, and has already been proven safe to turn on/off.

✕ **slackbot** dismissed their stale review 3 years ago

CAB approver has provided a review. Dismissing

🏷 **gsanford** added `cab-dep-ff-removal` `cab-dep-high-blast-radius` `cab-dep-standard-risk` labels on May 25, 2021

Logistics

Management 🤠

- Accountability
- Understand the risk profile of changes being merged

- **Weekly review** (GH tags), incidents, emergency merges, **Code Climate** velocity data, etc.

- Periodically **invite folks** to talk through **CAB-approved changes** that were later linked to **incidents**

Logistics

Review Date	PRs merged to chef-repo	Risk level		Blast radius level		Feature flags		
		Low risk	Standard risk	Low blast-radius	High blast-radius	Rollouts	New flags	Removed flags
05/07/2021	224	106	52	132	24	18	2	12
05/13/2021	241	111	52	132	30	46	14	15
05/21/2021	354	158	53	167	44	35	6	5
05/28/2021	303	171	44	172	38	34	9	8
06/04/2021	259	142	49	159	33	20	10	6
06/11/2021	324	167	53	188	32	32	10	5
06/18/2021	297	152	76	181	47	45	7	8
06/24/2021	328	184	70	191	64	47	10	10
07/02/2021	284	151	72	182	41	17	9	11
07/09/2021	303	195	59	216	38	44	10	5
07/15/2021	374	218	76	248	46	38	9	8
07/23/2021	315	166	77	189	55	31	5	9
07/30/2021	234	218	64	234	49	52	3	2

Logistics

Code Climate				Incidents			
PR Throughput	Time to Review	Revert Rate	Time to Merge	Emergency Merges	Failed Changes	# of (potentially) chef-related Incidents	# of (actually) chef-related Incidents
trending up	trending down	trending stable			0		
trending up	trending down	trending stable		2		1	
trending up	trending down	trending up		3		6	0
trending up	trending up	trending down		1	0	3	2
trending up	trending down	trending stable	trending down	0		1	1
trending up	trending stable; ⟨	trending down	trending down	2	1	1	1
trending up	trending down	trending stable	trending down	1	0	1	1
trending up	trending down	trending slightly down	trending down	3	0	4	0
trending stable; ⟨	trending down	trending up	trending down	1	1	4	1
trending slightly ⟨	trending down	trending stable	trending down	0	0	3	0
trending slightly ⟨	trending down	trending up	trending down	0	0	2	1
trending slightly ⟨	trending down	trending stable	trending down	1	0	4	1
trending stable	trending down	trending stable	trending down	3	0	3	1

Rollout

- Held multiple live trainings across multiple time zones, starting mid-April 2021.

- We first trained non-CAB approver chef experts. Incorporated feedback on the training process.

- Within a month, we'd trained > 20 engineers (tripling the number of CAB approvers) across nearly all infrastructure teams operating in `chef-repo` regularly.

⚲ 6 élément(s) épinglé(s)

Low Risk Change FLUX DE TRAVAIL 12 h (27 mai 2021 ∨   ← **Six weeks later**

⚪ @msg is requesting approval for a low-risk change.

🛡️ 1   ✅ 1   ☺️⁺

⑂ 👤 **5 réponses** Dernière réponse il y a 3 ans

Maude 12 h 45

@msg I think you'll have the last ever #infra-change-control entry!

V Brennan 💬 12 h 46

I might cry!!!! @maude you da best ❤️ 👧 ER

💖 1   🕷️ 1   ✴️ 2   ☺️⁺

Maude 12 h 46

FYI We're archiving this channel!

👤 **Maude**

🦊 We're deprecating #infra-change-control!

Between the new labels added as part of the CAB deputy process and #chef-deploys, we're capturing all the information we need to understand what kinds of changes are making it out the door.

Later today, we'll archive #infra-change-control. Updates to the training doc are already live to reflect that authors no longer need to submit the workflow and deputies no longer need to approve in thread.

Everything will be done in GitHub 🐙 from here on out!
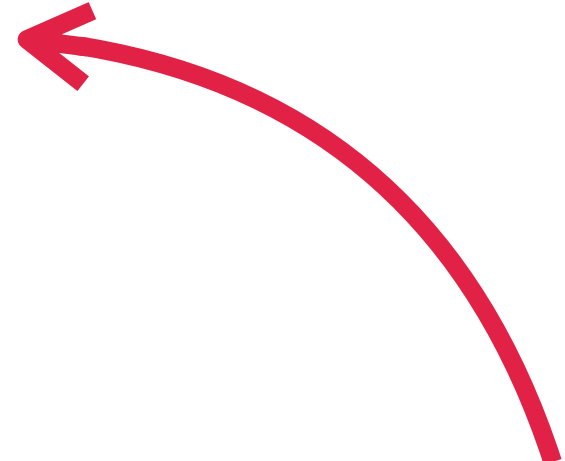
… and just a few months later

07/15/2021	374	218	76	248	46	38	9	8
07/23/2021	315	166	77	189	55	31	5	9
07/30/2021	234	218	64	234	49	52	3	2

- After 2 months, in July, we dropped the weekly metrics sync, moving to once a month.

- After another 2 months, in September, we dropped the meeting entirely.

- We'd also made significant progress on the goals set in #infrastructure-safety!

🛡️ CAB Deputies **Today**

📓 **234 quiz submissions**
✅ **220 deputies**

… and growing!

	chef-repo	
Distinct contributor count (Feb 2024 – March 2024)	209	
Total PR count (Feb 2024 – March 2024)	2,508	

Merci 👋

You can find me most
places on the internet as
@qcmaude