Next Generation of CI/CD: Analytics-driven Traffic Management on Kubernetes

Fabio A. Oliveira, IBM Research

iter8.tools

Teamwork

iter8 project

- https://iter8.tools
- https://github.com/iter8-tools

iter8 team

- Alan Cha
- Fabio Oliveira
- Hai Huang
- Michael Kalantar
- Srinivasan Parthasarathy
- Sushma Ravichandran
- Yew Huey Liu
- Zijun Wang

- Collaborator
  - Lionel Villard

What to expect

- This is **NOT**
  - yet another canary release demo on K8s
  - yet another spin on how to integrate tools for canary releases

What to expect

- This is **NOT**
  - yet another canary release demo on K8s
  - yet another spin on how to integrate tools for canary releases

- What's the goal then?
  - Raise awareness of a **fundamental yet largely ignored** problem at the core of cloud native canary releases, performance tests, and A/B & A/B/n testing
  - Offer an open solution to that problem and engage the community
  - And yes, there will be demos… ☺

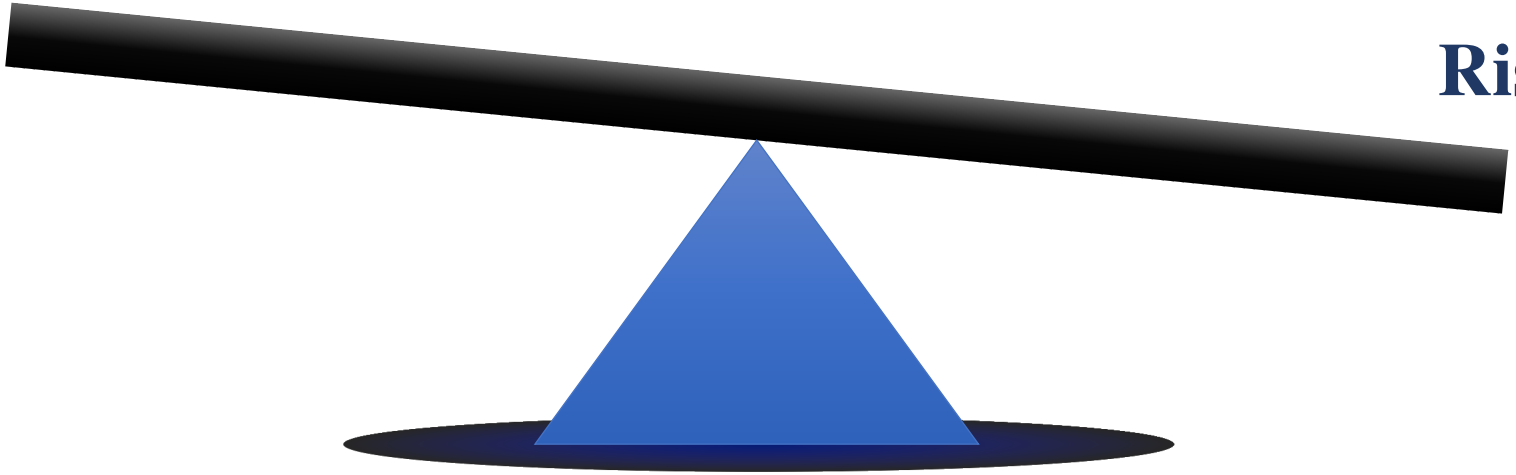
It is about agility, but…

Attract new clients
Stay relevant in the market
New features and bug fixes

Speed & Agility

SLO violations
Bugs
Reputation
Decline in revenue

Risks



… it is also about learning…

- What if you could *safely*
  - learn how your code behaves in production or test?

- What if you could *continuously* and *safely*
  - learn what resonates with your users?
  - find ways to increase your company's revenue?
  - maximize your company's revenue *as you learn*?

… it is also about learning…

- What if you could *safely*
  - learn how your code behaves in production or test?

performance test
canary releases

- What if you could *continuously* and *safely*
  - learn what resonates with your users?
  - find ways to increase your company's revenue?
  - maximize your company's revenue *as you learn*?

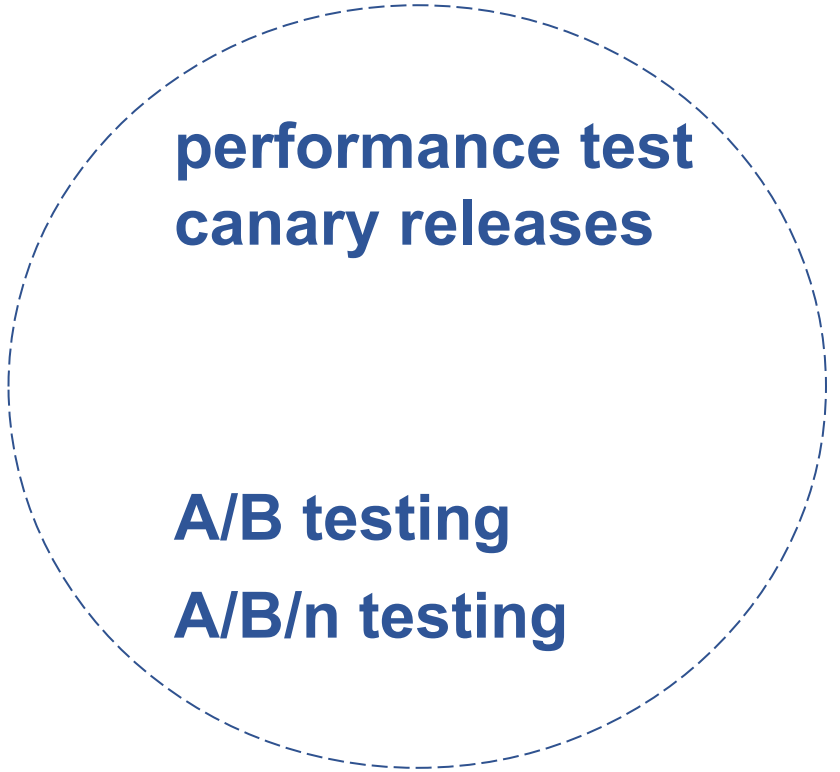
A/B testing

A/B/n testing

… it is also about learning…

- What if you could *safely*
  - learn how your code behaves in production or test?

- What if you could *continuously* and *safely*
  - learn what resonates with your users?
  - find ways to increase your company's revenue?
  - maximize your company's revenue *as you learn*?



… it is also about learning…

- What if you could *safely*
  - learn how your code behaves in production or test?

- What if you could *continuously* and *safely*
  - learn what resonates with your users?
  - find ways to increase your company's revenue?
  - maximize your company's revenue *as you learn*?

performance test
canary releases

A/B testing

A/B/n testing

How to unleash the power of ***continuous experimentation*** in K8s?

… and analytics is crucial

Make no mistake, continuous experimentation is an **analytics problem**!

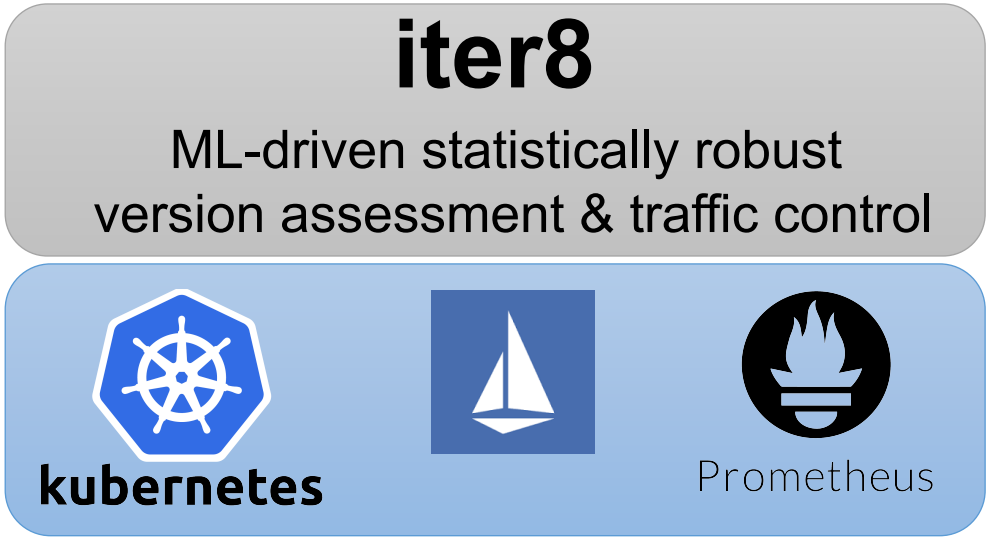
And a ***comparative*** **analytics problem**!

… and analytics is crucial

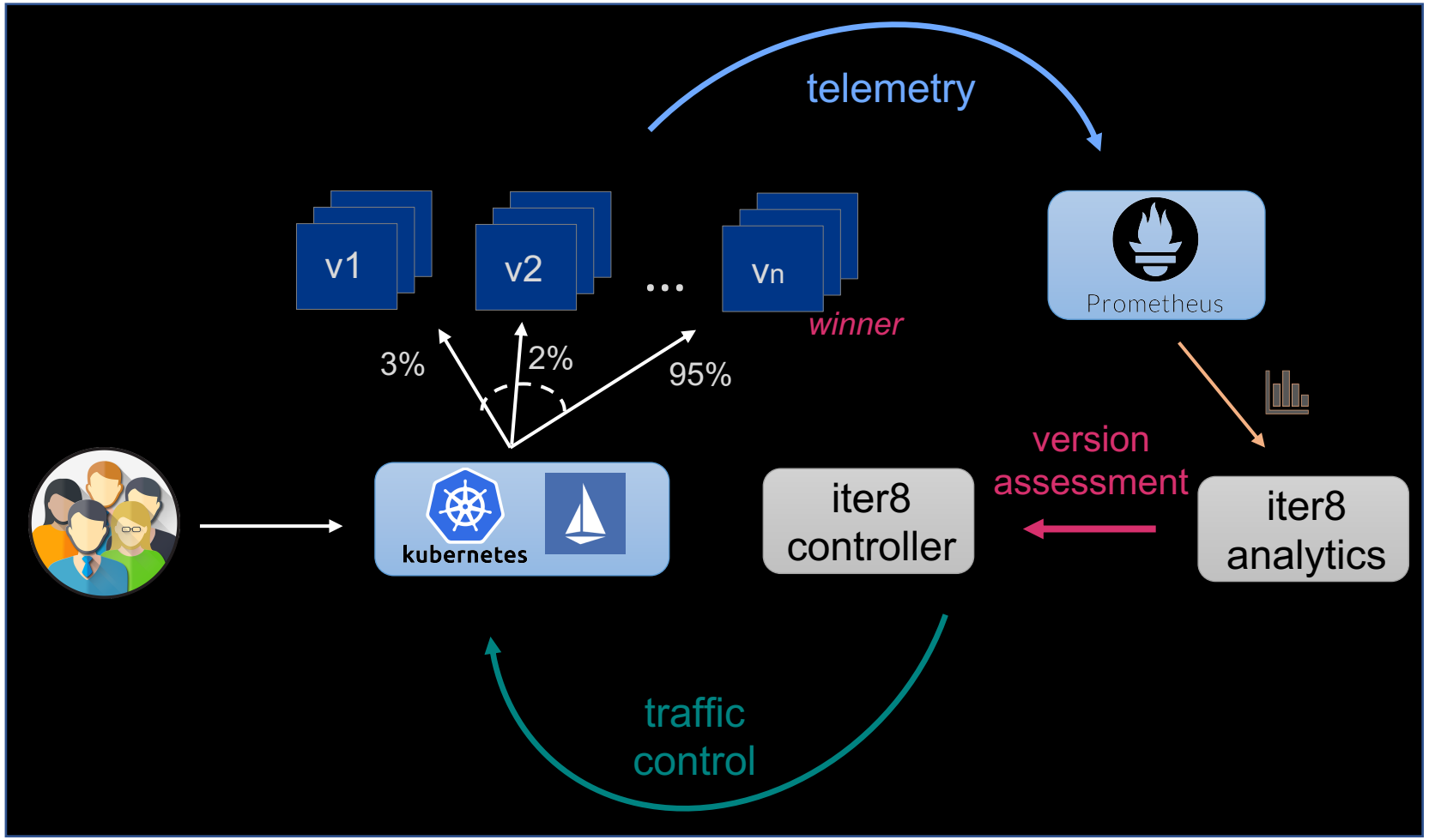
Make no mistake, continuous experimentation is an **analytics problem**!

And a ***comparative*** **analytics problem**!

Enter iter8…



iter8 overview



- Version assessment
  - With confidence
- Traffic control strategies
  - Progressive
  - Top-2
  - Uniform
- Traffic control safety filters
  - Cutoff on failure
  - Maximum increment
  - Match clause
  - Experiment traffic percentage
- iter8-trend (another component)
  - Uncover trends as a service evolves

iter8 experiment types

- Performance test
  - 1 version
  - Assess version against criteria
  - Typically done in a test/dev environment
  - Can be done in production, too

```yaml
apiVersion: iter8.tools/v1alpha2
kind: Experiment
metadata:
  name: my-perf-test
spec:
  service:
    name: my-service
    baseline: my-service-v1
  criteria:
    - metric: iter8_mean_latency
      threshold:
        type: absolute
        value: 120
    - metric: iter8_error_rate
      threshold:
        type: absolute
        value: 0.0001
  duration:
    interval: 20s
    maxIterations: 20
```

iter8 experiment types

- Canary release
  - 2 versions: *baseline* & *candidate*
  - Assess canary
    - make sure no SLOs are violated
    - ***relative criteria*** might make sense
  - Apply traffic control strategy
  - If canary passes, roll forward
  - If canary fails, roll back

```yaml
apiVersion: iter8.tools/v1alpha2
kind: Experiment
metadata:
  name: my-canary-test
spec:
  service:
    name: my-service
    baseline: my-service-v1
    candidates:
      - my-service-v2
  criteria:
    - metric: iter8_mean_latency
      threshold:
        type: absolute
        value: 200
    - metric: iter8_error_rate
      threshold:
        type: relative
        value: 1.05
  duration:
    interval: 20s
    maxIterations: 20
  trafficControl:
    strategy: progressive
    maxIncrement: 10
```

iter8 experiment types

- Canary release
  - 2 versions: *baseline* & *candidate*
  - Assess canary
    - make sure no SLOs are violated
    - ***relative criteria*** might make sense
  - Apply traffic control strategy
  - If canary passes, roll forward
  - If canary fails, roll back

```yaml
apiVersion: iter8.tools/v1alpha2
kind: Experiment
metadata:
  name: my-canary-test
spec:
  service:
    name: my-service
    baseline: my-service-v1
    candidates:
      - my-service-v2
  criteria:
    - metric: iter8_mean_latency
      threshold:
        type: absolute
        value: 200
    - metric: iter8_error_rate
      threshold:
        type: relative
        value: 1.05
  duration:
    interval: 20s
    maxIterations: 20
  trafficControl:
    strategy: progressive
    maxIncrement: 10
```

iter8 experiment types

- Canary release
  - 2 versions: *baseline* & *candidate*
  - Assess canary
    - make sure no SLOs are violated
    - ***relative criteria*** might make sense
  - Apply traffic control strategy
  - If canary passes, roll forward
  - If canary fails, roll back

```yaml
apiVersion: iter8.tools/v1alpha2
kind: Experiment
metadata:
  name: my-canary-test
spec:
  service:
    name: my-service
    baseline: my-service-v1
    candidates:
      - my-service-v2
  criteria:
    - metric: iter8_mean_latency
      threshold:
        type: absolute
        value: 200
    - metric: iter8_error_rate
      threshold:
        type: relative
        value: 1.05
  duration:
    interval: 20s
    maxIterations: 20
  trafficControl:
    strategy: progressive
    maxIncrement: 10
```

iter8 experiment types

- A/B & A/B/n testing
  - "n" versions:
    - 1 *baseline*
    - 1 or more *candidates*
  - Compare versions to declare a winner
    - maximize a *reward* metric
    - AND make sure no SLOs are violated
  - Apply traffic control strategy
  - Traffic will shift towards winner

```yaml
metadata:
  name: my-abn-test
spec:
  service:
    name: my-service
    baseline: my-service-v1
    candidates:
      - my-service-v2
      - my-service-v3
  criteria:
    - metric: iter8_mean_latency
      threshold:
        type: absolute
        value: 200
    - metric: iter8_error_rate
      threshold:
        type: relative
        value: 1.05
    - metric: mean_books_purchased
      isReward: true
  duration:
    interval: 20s
    maxIterations: 20
  trafficControl:
    strategy: progressive
    maxIncrement: 10
```

iter8 experiment types

- A/B & A/B/n testing
  - "n" versions:
    - 1 *baseline*
    - 1 or more *candidates*
  - Compare versions to declare a winner
    - maximize a *reward* metric
    - AND make sure no SLOs are violated
  - Apply traffic control strategy
  - Traffic will shift towards winner

```yaml
metadata:
  name: my-abn-test
spec:
  service:
    name: my-service
    baseline: my-service-v1
    candidates:
      - my-service-v2
      - my-service-v3
  criteria:
    - metric: iter8_mean_latency
      threshold:
        type: absolute
        value: 200
    - metric: iter8_error_rate
      threshold:
        type: relative
        value: 1.05
    - metric: mean_books_purchased
      isReward: true
  duration:
    interval: 20s
    maxIterations: 20
  trafficControl:
    strategy: progressive
    maxIncrement: 10
```

iter8 experiment types

- A/B & A/B/n testing
  - "n" versions:
    - 1 *baseline*
    - 1 or more *candidates*
  - Compare versions to declare a winner
    - maximize a *reward* metric
    - AND make sure no SLOs are violated
  - Apply traffic control strategy
  - Traffic will shift towards winner

```yaml
metadata:
  name: my-abn-test
spec:
  service:
    name: my-service
    baseline: my-service-v1
    candidates:
      - my-service-v2
      - my-service-v3
  criteria:
    - metric: iter8_mean_latency
      threshold:
        type: absolute
        value: 200
    - metric: iter8_error_rate
      threshold:
        type: relative
        value: 1.05
    - metric: mean_books_purchased
      isReward: true
  duration:
    interval: 20s
    maxIterations: 20
  trafficControl:
    strategy: progressive
    maxIncrement: 10
```

Demo 1

- Automated A/B/n testing
  - 3 competing versions
  - Goal
    - Maximize the value of a business-oriented metric…
      - … while satisfying SLOs

DEMO…

Demo 2

- ***Human-in-the-loop*** A/B/n testing
  - Same scenario as before
  - This time, the user is in control
    - Gets insights / assessments
    - Decides what to do
- Show integration between iter8 & KUI
  - KUI – visual terminal
    - https://kui.tools

DEMO…

Summary

- iter8's goals
  - Unleash the power of continuous experimentation
    - ML-driven statistically robust
      - version assessment
      - traffic control
  - Uncover trends as a service evolves
    - iter8-trend
- Current support
  - Kubernetes
  - Istio (control)
  - Prometheus (metrics)
- Additional integrations
  - Kiali
    - Istio's open source UI
      - https://kiali.io
    - Kiali's iter8 extension
  - KUI
    - Visual terminal
      - https://kui.tools

Where to find us

KubeCon CloudNativeCon Europe 2020 Virtual

iter8 project

- https://iter8.tools
- https://github.com/iter8-tools

Thank you!

KubeCon
CloudNativeCon
Europe 2020
Virtual
HELM
KEEP CLOUD NATIVE
CONNECTED