



KubeCon



CloudNativeCon

North America 2019

Leveling Up Your CD: Unlocking Progressive Delivery on Kubernetes

Danny Thomson & Jesse Suen, Intuit



Who are we

intuit.



KubeCon



CloudNativeCon

North America 2019

1983

Founded



1993

IPO



5,000

Developers



20

Locations



\$6.8B
FY19

Revenue



50M

Customers



Some Intuit Statistics



KubeCon



CloudNativeCon

North America 2019

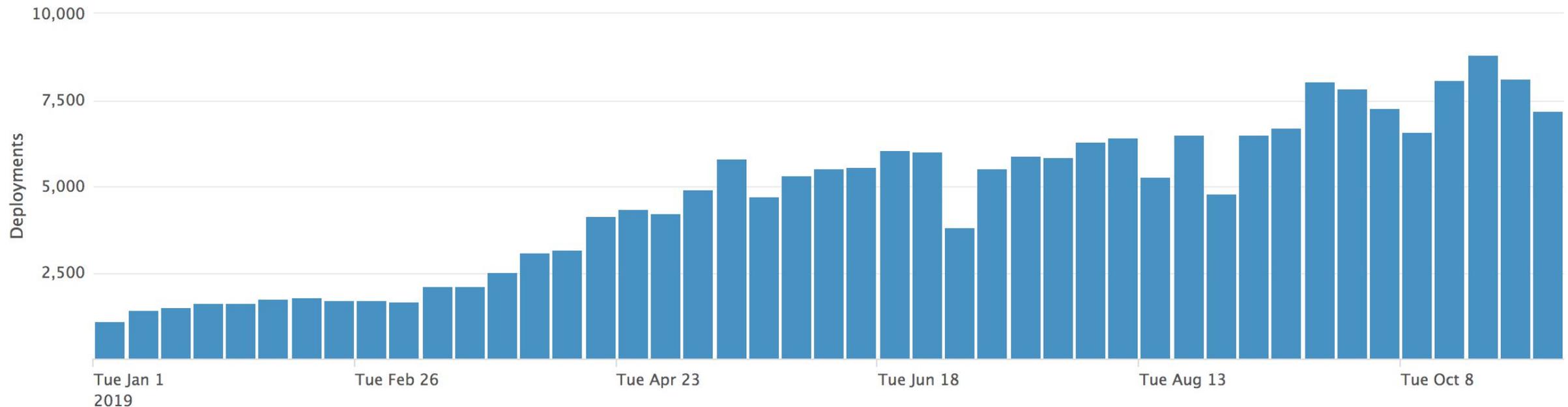
intuit.

- 4 business units
- 30 business segments
- 1,200+ developers using Kubernetes



kubernetes

- 160+ clusters (Intuit managed)
- 6,600 nodes
- 5,400 namespaces
- 62,000 pods
- **1,300 deploys a day**



Continuous Delivery



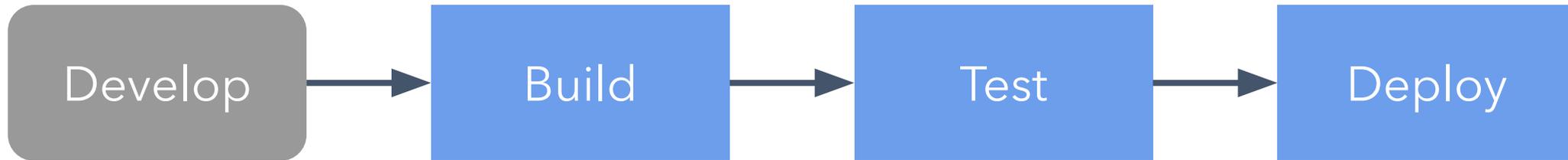
KubeCon



CloudNativeCon

North America 2019

- Shortens the time it takes to deliver software to users



- Delivers faster, but not necessarily safer
- Most problems and outages occur after a change



Progressive Delivery

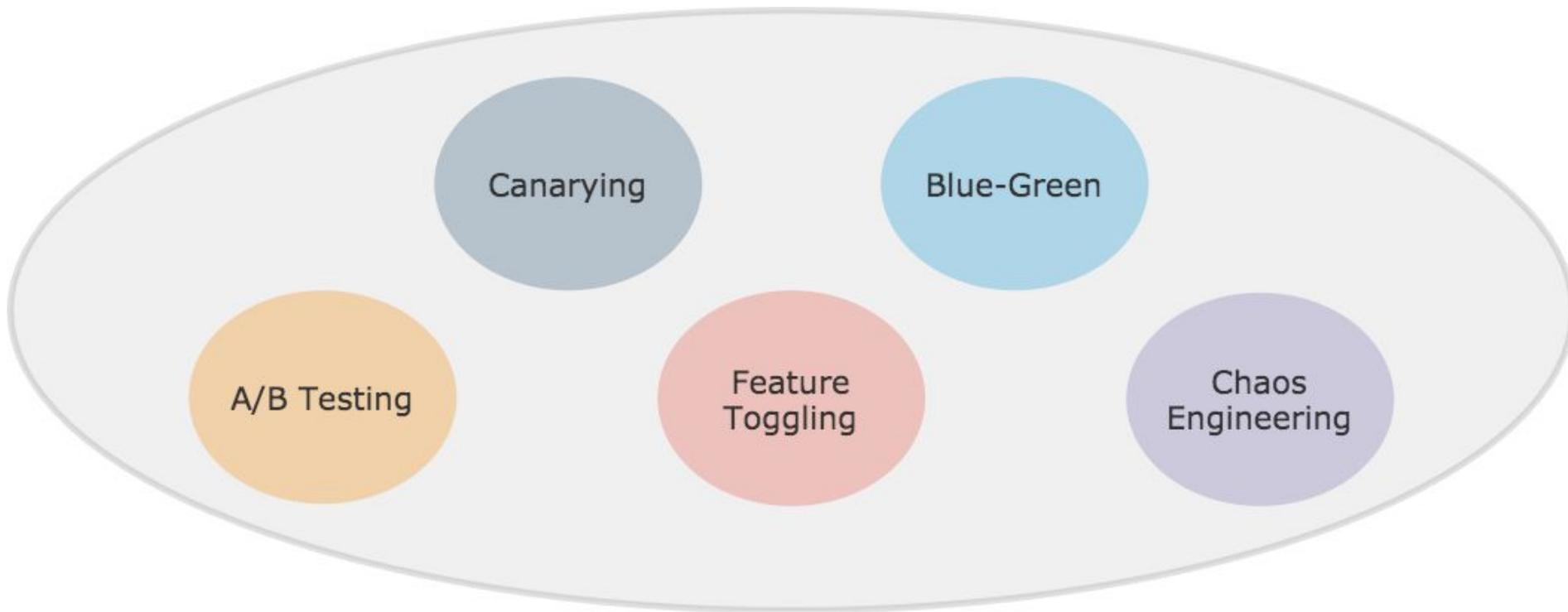


KubeCon



CloudNativeCon

North America 2019



“Progressive delivery is continuous delivery with fine-grained control over the blast radius.”

— James Governor, RedMonk



Progressive Delivery

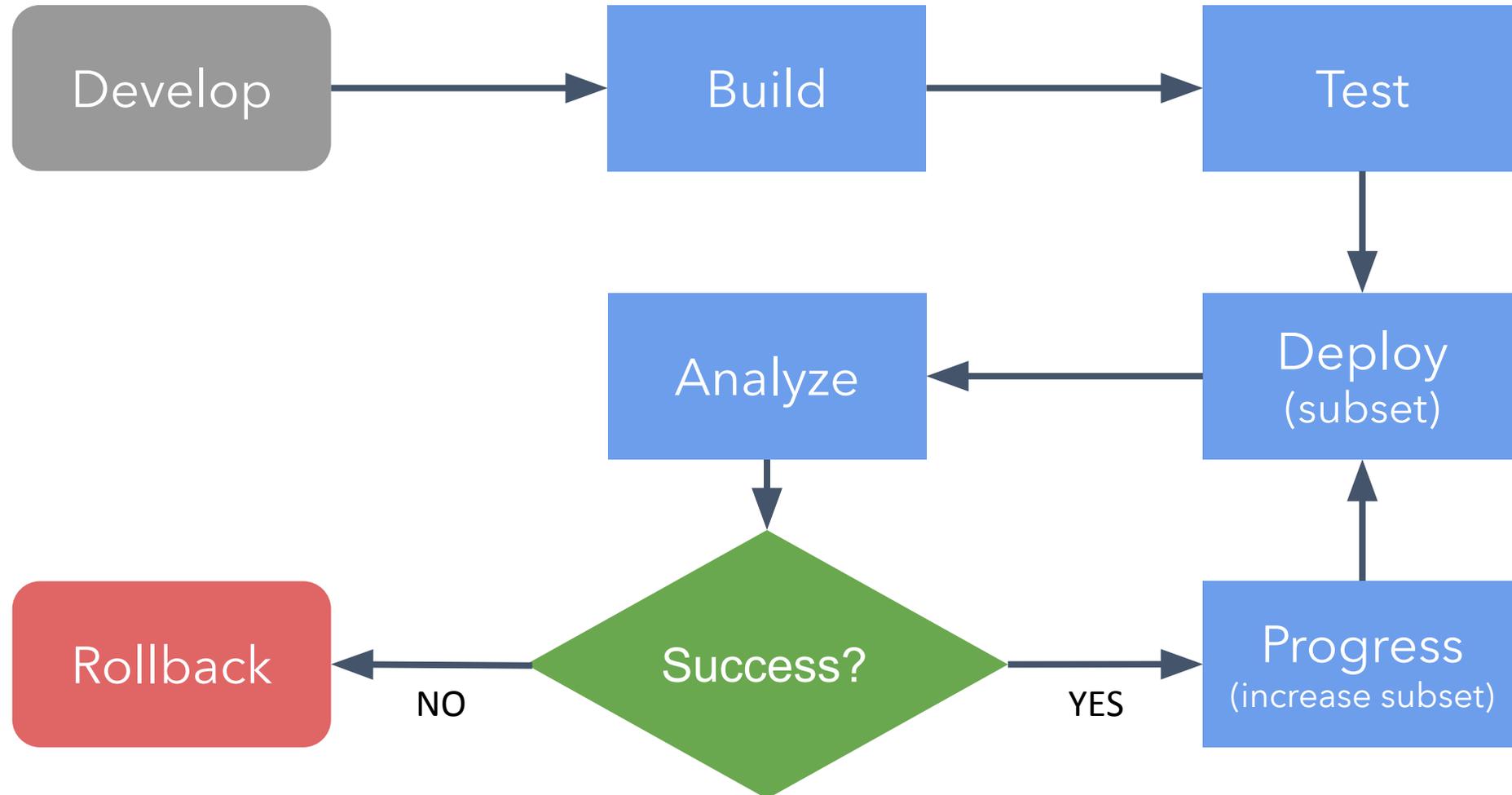


KubeCon



CloudNativeCon

North America 2019



Progressive Delivery



KubeCon



CloudNativeCon

North America 2019

User Segmentation



Traffic Percentage
Headers & Cookies
Geographical

Traffic Management



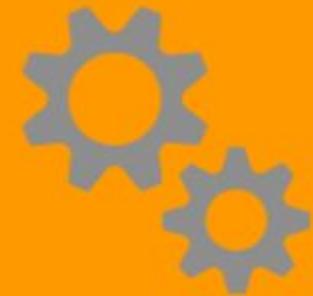
Service Mesh
Ingress Controllers

Observability



Tracing
Logging
Metrics

Automation



Promotion
Rollback



Problem



KubeCon



CloudNativeCon

North America 2019

- Rolling Update provides few controls over speed
- Container readiness probes are not enough
 - Unsuitable for deeper or temporary checks
 - Unable to use external metrics
- Able to halt the progression, but not reverse



Use Cases



KubeCon



CloudNativeCon

North America 2019

How do I...

- automatically rollback an update due to failed metrics
- fine-tune my success and failure criteria
- insert a manual judgement step
- use my own business metrics for analysis
- experiment with multiple versions of my service (e.g. baseline vs. canary, A/B testing)
- and others...



Requirements



KubeCon



CloudNativeCon

North America 2019

01

Robust

- Does not rely on scripting/pipelines

02

Standardized

- Use industry standard tooling (Prometheus, Kayenta, Wavefront, etc...)

03

Flexible

- Control over the rollout plan
- What metrics to analyze and from where

04

Declarative

- GitOps focused



Argo Rollouts



KubeCon



CloudNativeCon

North America 2019

Phase 1: Deployment++

- Drop-in replacement for a Deployment
- Additional deployment strategies: blue-green and canary
- Declarative and GitOps friendly



Rollout



KubeCon



CloudNativeCon

North America 2019

```
apiVersion: argoproj.io/v1alpha1
kind: Rollout
metadata:
  name: canary-demo
spec:
  replicas: 5
  template:
    spec:
      containers:
      - name: app
        image: argoproj/rollouts-demo:blue
    ...
  strategy:
    canary:
      steps:
      - setWeight: 40
      - pause: {duration: 3600}
      - setWeight: 60
      - pause: {duration: 10}
      - setWeight: 80
      - pause: {duration: 10}
```

- Manages creation, scaling, and deletions of ReplicaSets



Rollout



KubeCon



CloudNativeCon

North America 2019

```
apiVersion: argoproj.io/v1alpha1
kind: Rollout
metadata:
  name: canary-demo
spec:
  replicas: 5
  template:
    spec:
      containers:
      - name: app
        image: argoproj/rollouts-demo:blue
    ...
  strategy:
    canary:
      steps:
      - setWeight: 40
      - pause: {duration: 3600}
      - setWeight: 60
      - pause: {duration: 10}
      - setWeight: 80
      - pause: {duration: 10}
```

- Spec is mostly identical to Deployment



Rollout



KubeCon



CloudNativeCon

North America 2019

```
apiVersion: argoproj.io/v1alpha1
kind: Rollout
metadata:
  name: canary-demo
spec:
  replicas: 5
  template:
    spec:
      containers:
      - name: app
        image: argoproj/rollouts-demo:blue
    ...
  strategy:
    canary:
      steps:
      - setWeight: 40
      - pause: {duration: 3600}
      - setWeight: 60
      - pause: {duration: 10}
      - setWeight: 80
      - pause: {duration: 10}
```

- New blue-green and canary strategies provides control over *how* to update the stable version to new version



Argo Rollouts



KubeCon



CloudNativeCon

North America 2019

Phase 1: Deployment++

- Drop-in replacement for a Deployment
- Additional deployment strategies: blue-green and canary
- Declarative and GitOps friendly

Phase 2: Progressive Delivery

- Analysis
- Experimentation



Analysis CRDs



KubeCon



CloudNativeCon

North America 2019

- Brings observability to the delivery process
- Defines *how* to perform a canary analysis:
 - What **metrics** to measure and **when**
 - What **values** are considered successful, failed, inconclusive
- Automates promotion & rollback



Rollout Integration



KubeCon



CloudNativeCon

North America 2019

```
apiVersion: argoproj.io/v1alpha1
kind: Rollout
metadata:
  name: canary-demo
spec:
  replicas: 5
  template:
    spec:
      containers:
      - name: app
        image: argoproj/rollouts-demo:blue
    ...
  strategy:
    canary:
      analysis:
        templateName: success-rate
      steps:
      - setWeight: 40
      - pause: {duration: 3600}
      - setWeight: 60
      - pause: {duration: 10}
      - setWeight: 80
      - pause: {duration: 10}
```

Canary Analysis

- Analysis is performed in the background, while the rollout is progressing through its steps
- Started at the beginning of a rollout, and stopped when the rollout is complete



Rollout Integration



KubeCon



CloudNativeCon

North America 2019

```
apiVersion: argoproj.io/v1alpha1
kind: Rollout
metadata:
  name: canary-demo
spec:
  replicas: 5
  template:
    spec:
      containers:
      - name: app
        image: argoproj/rollouts-demo:blue
    ...
  strategy:
    canary:
      steps:
      - setWeight: 20
      - analysis:
          templateName: http-benchmark
      - setWeight: 40
```

Inline Analysis

- Analysis can also be performed “inline,” as a blocking step in the rollout
- Suitable for more heavyweight analysis where recurrence may not be desired (e.g. benchmarking, load/stress testing, integration tests)





KubeCon



CloudNativeCon

North America 2019

DEMO



AnalysisTemplate



KubeCon



CloudNativeCon

North America 2019

```
apiVersion: argoproj.io/v1alpha1
kind: AnalysisTemplate
metadata:
  name: success-rate
spec:
  args:
  - name: ingress
  metrics:
  - name: success-rate
    interval: 5m
    count: 5
    successCondition: result[0] > 0.90
    failureLimit: 2
    provider:
      prometheus:
        address: http://prometheus-svc.prometheus-ns:9090
        query: >-
          sum(rate/nginx_ingress_controller_requests
            {ingress="{args.ingress}"status!~"[4-5].*" } [5m]))
          /
          sum(rate/nginx_ingress_controller_requests
            {ingress="{args.ingress}" } [5m]))
```

Defines one or more key metrics to monitor during a rollout

Support for many providers:

- Prometheus
- Job
- Kayenta
- Web (coming)
- Wavefront (coming)
- and others..



AnalysisTemplate - Prometheus Provider



KubeCon



CloudNativeCon

North America 2019

```
apiVersion: argoproj.io/v1alpha1
kind: AnalysisTemplate
metadata:
  name: success-rate
spec:
  args:
  - name: ingress
  metrics:
  - name: success-rate
    interval: 5m
    count: 5
    successCondition: result[0] > 0.90
    failureLimit: 2
  provider:
    prometheus:
      address: http://prometheus-svc.prometheus-ns:9090
      query: >-
        sum(rate(nginx_ingress_controller_requests
          {ingress="{args.ingress}"status!~"[4-5].*"} [5m]))
        /
        sum(rate(nginx_ingress_controller_requests
          {ingress="{args.ingress}"} [5m]))
```

Prometheus Provider

- Address - prometheus server
- Query - PromQL query

Example (HTTP success rate):

of non-4xx/5xx HTTP requests

of total HTTP requests



AnalysisTemplate - Success Condition



KubeCon



CloudNativeCon

North America 2019

```
apiVersion: argoproj.io/v1alpha1
kind: AnalysisTemplate
metadata:
  name: success-rate
spec:
  args:
  - name: ingress
  metrics:
  - name: success-rate
    interval: 5m
    count: 5
    successCondition: result[0] > 0.90
    failureLimit: 2
  provider:
    prometheus:
      address: http://prometheus-svc.prometheus-ns:9090
      query: >-
        sum(rate(nginx_ingress_controller_requests
          {ingress="{args.ingress}"status!~"[4-5].*"}[5m]))
        /
        sum(rate(nginx_ingress_controller_requests
          {ingress="{args.ingress}"}[5m]))
```

- An **expression** which interprets the result of a measurement
- Results can return as:
 - scalars
 - vectors
 - structured objects
- Built-in functions like `any()`, `all()`, `filter()`, `map()`
- Results can also be **Inconclusive** to allow for manual judgements



AnalysisTemplate - Interval & Count



KubeCon



CloudNativeCon

North America 2019

```
apiVersion: argoproj.io/v1alpha1
kind: AnalysisTemplate
metadata:
  name: success-rate
spec:
  args:
  - name: ingress
  metrics:
  - name: success-rate
    interval: 5m
    count: 5
    successCondition: result[0] > 0.90
    failureLimit: 2
  provider:
    prometheus:
      address: http://prometheus-svc.prometheus-ns:9090
      query: >-
        sum(rate(nginx_ingress_controller_requests
          {ingress="{args.ingress}"status!~"[4-5].*"} [5m]))
        /
        sum(rate(nginx_ingress_controller_requests
          {ingress="{args.ingress}"} [5m]))
```

- **Interval**
 - How frequent to query the provider
- **Count**
 - Number of times to take a measurement
 - Runs indefinitely if omitted (or until failure)



AnalysisTemplate - Arguments



KubeCon



CloudNativeCon

North America 2019

```
apiVersion: argoproj.io/v1alpha1
kind: AnalysisTemplate
metadata:
  name: success-rate
spec:
  args:
  - name: ingress
  metrics:
  - name: success-rate
    interval: 5m
    count: 5
    successCondition: result[0] > 0.90
    failureLimit: 2
    provider:
      prometheus:
        address: http://prometheus-svc.prometheus-ns:9090
        query: >-
          sum(rate/nginx_ingress_controller_requests
            {ingress={{args.ingress}}status!~"[4-5].*" } [5m]))
          /
          sum(rate/nginx_ingress_controller_requests
            {ingress={{args.ingress}} } [5m]))
```

- **Arguments** make Analysis Templates parameterizable
- Enables templates to be reusable/standardized across organizations and communities
- Makes templates building blocks for higher levels resources



Experiment CRD



KubeCon



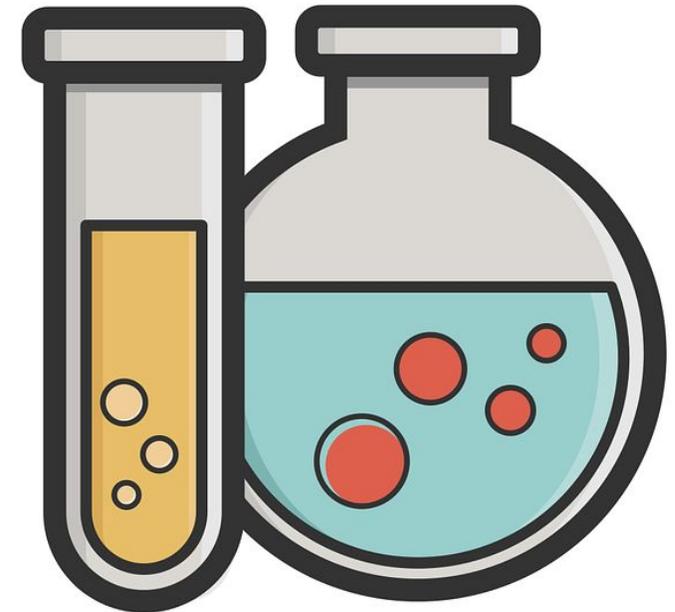
CloudNativeCon

North America 2019

- An **ephemeral** run of one or more versions of a service
- Coupled with analysis
- Can be started as a Rollout step

Use Cases:

- A/B testing
- Baseline vs. Canary Analysis (Kayenta)
- ML model testing



Experiment CRD



KubeCon



CloudNativeCon

North America 2019

```
apiVersion: argoproj.io/v1alpha1
kind: Experiment
metadata:
  name: demo-ab-test
spec:
  duration: 15m
  templates:
  - name: purple
    template:
      spec:
        containers:
        - name: rollouts-demo
          image: argoproj/rollouts-demo:purple
    ...
  - name: orange
    template:
      spec:
        containers:
        - name: rollouts-demo
          image: argoproj/rollouts-demo:orange
    ...
  analyses:
  - name: purple
    templateName: http-benchmark
    args: [{name: host, value: purple}]
  - name: orange
    templateName: http-benchmark
    args: [{name: host, value: orange}]
```



Experiment CRD



KubeCon



CloudNativeCon

North America 2019

```
apiVersion: argoproj.io/v1alpha1
kind: Experiment
metadata:
  name: demo-ab-test
spec:
  duration: 15m
  templates:
  - name: purple
    template:
      spec:
        containers:
        - name: rollouts-demo
          image: argoproj/rollouts-demo:purple
    ...
  - name: orange
    template:
      spec:
        containers:
        - name: rollouts-demo
          image: argoproj/rollouts-demo:orange
    ...
  analyses:
  - name: purple
    templateName: http-benchmark
    args: [{name: host, value: purple}]
  - name: orange
    templateName: http-benchmark
    args: [{name: host, value: orange}]
```

- Starts multiple versions of a service at the same time
- Runs for a specified duration (or indefinitely until failure)



Experiment CRD



KubeCon



CloudNativeCon

North America 2019

```
apiVersion: argoproj.io/v1alpha1
kind: Experiment
metadata:
  name: demo-ab-test
spec:
  duration: 15m
  templates:
  - name: purple
    template:
      spec:
        containers:
        - name: rollouts-demo
          image: argoproj/rollouts-demo:purple
        ...
  - name: orange
    template:
      spec:
        containers:
        - name: rollouts-demo
          image: argoproj/rollouts-demo:orange
        ...
  analyses:
  - name: purple
    templateName: http-benchmark
    args: [{name: host, value: purple}]
  - name: orange
    templateName: http-benchmark
    args: [{name: host, value: orange}]
```

- Coordinates analysis with the templates' readiness
- Can automatically shut down experiments which are not meeting metric requirements



AnalysisTemplate - Job Provider



KubeCon



CloudNativeCon

North America 2019

```
apiVersion: argoproj.io/v1alpha1
kind: AnalysisTemplate
metadata:
  name: http-benchmark
spec:
  args:
  - name: host
  metrics:
  - name: http-benchmark
    provider:
      job:
        spec:
          template:
            spec:
              containers:
              - name: load-tester
                image: argoproj/load-tester:latest
                command: [sh, -xec]
                args:
                - |
                    wrk -t1 -c1 -d10s -s report.lua \
                      http://{{args.host}}/color
                    jq -e '.errors_ratio <= 0.05' report.json
```

Job based metric

- Exit code determines success or failure

Example:

- Runs a http benchmark against a supplied host
- Verify error rate is $\leq 5\%$





KubeCon



CloudNativeCon

North America 2019

DEMO



Summary



KubeCon



CloudNativeCon

North America 2019

- Analysis and Experiments are building blocks
- Customize to your progressive delivery plan
- Progressive Delivery is multi-faceted
- Just the beginning!



What's Next



KubeCon



CloudNativeCon

North America 2019

- Service Mesh & Ingress Controller integration
 - Finer grained traffic shaping
 - Advanced user segmentation
- Additional metric providers
 - Contributions welcome!



Links



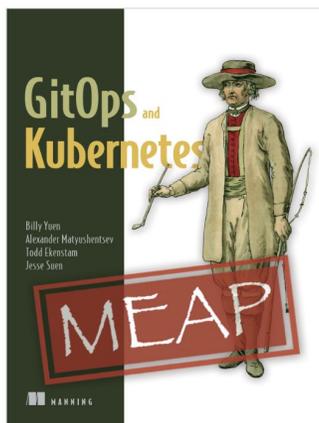
KubeCon



CloudNativeCon

North America 2019

- Argo Rollouts: <https://github.com/argoproj/argo-rollouts>
- Demo: <https://github.com/argoproj/rollouts-demo>
- Come find us at the Intuit booth S47



50% Code: **mlyuen**



<http://bit.ly/gitops-and-k8s>



SCAN ME

