

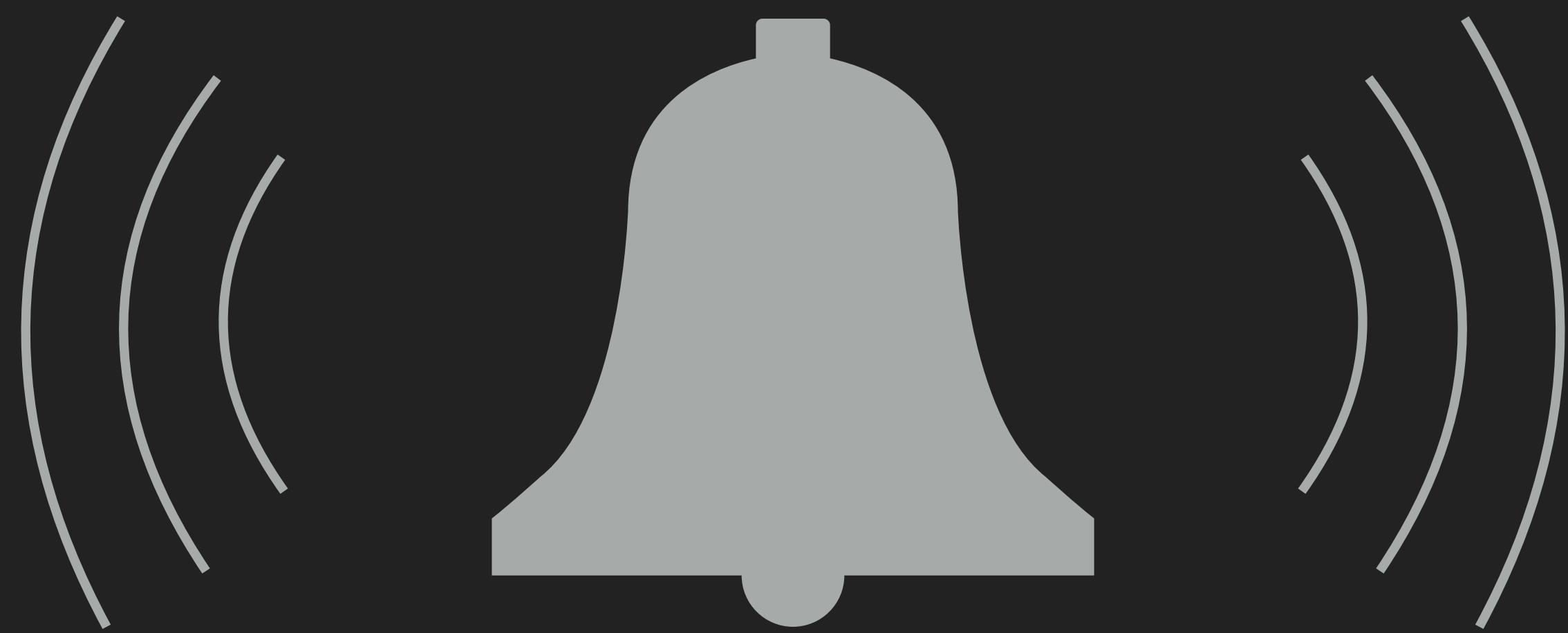


GET IN THE GAME: How to Use Referee and Kayenta to Do Canary Analysis

Justin Field and Melana Hammel

Nike









Canary deployments introduce a
code change to a small percentage
of your traffic



VALUE OF CANARY DEPLOYMENTS



Catch errors



Limit blast radius



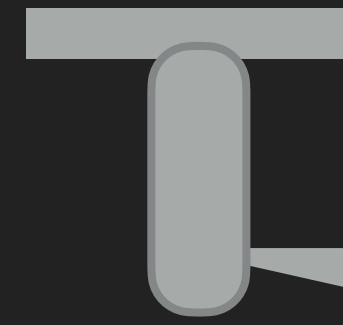
Save engineers time



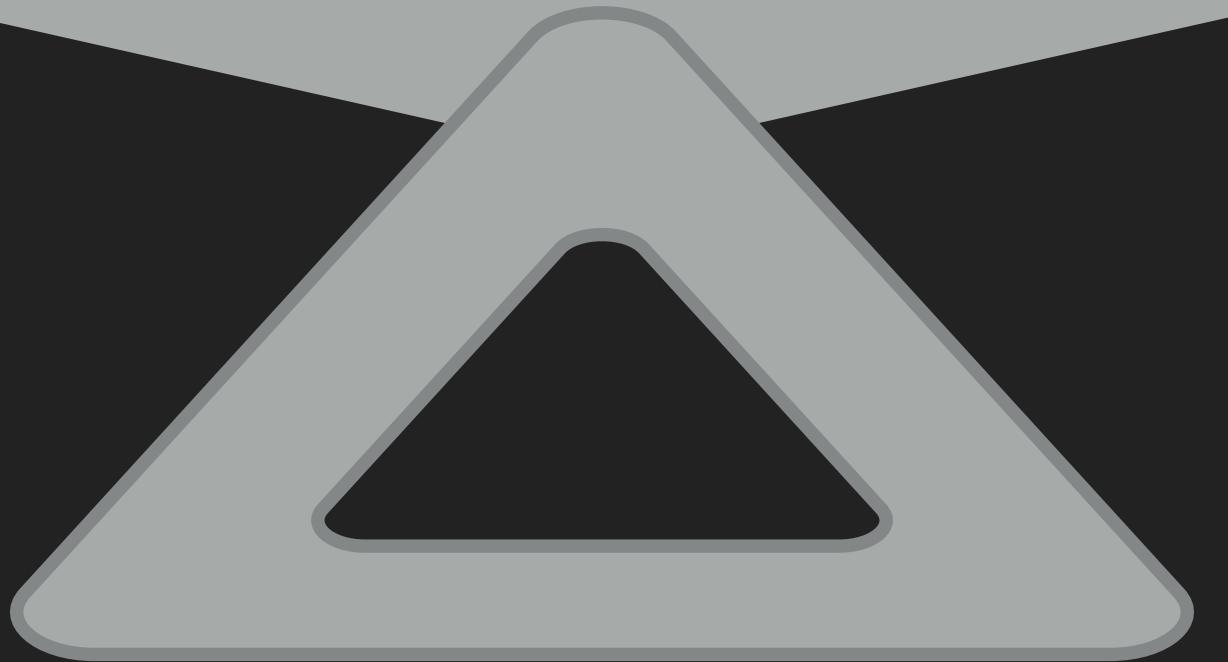
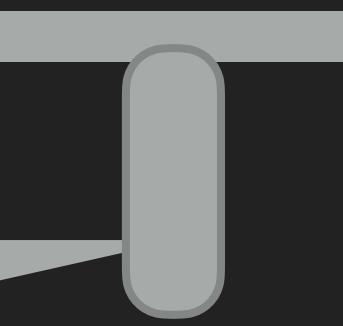
We needed the ability to canary at Nike



BUILD



BUY





Spinnaker



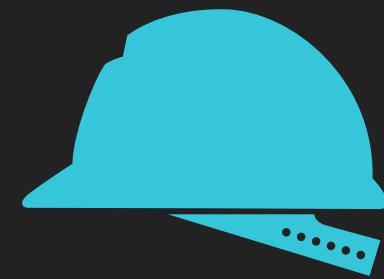
Halyard



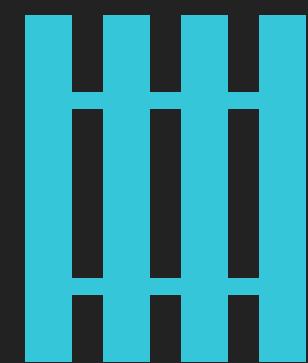
Clouddriver



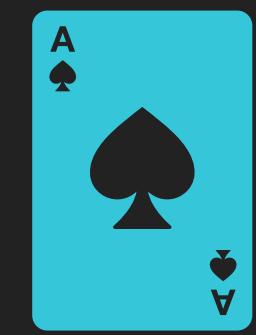
Echo



Igor



Gate



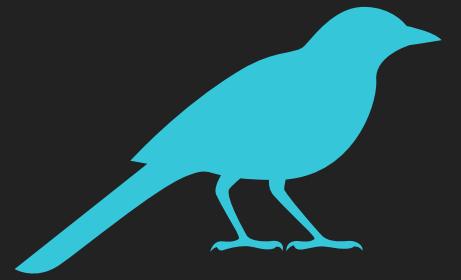
Deck



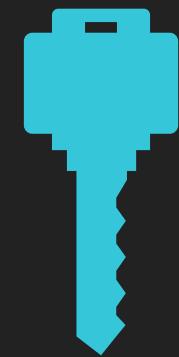
Front50



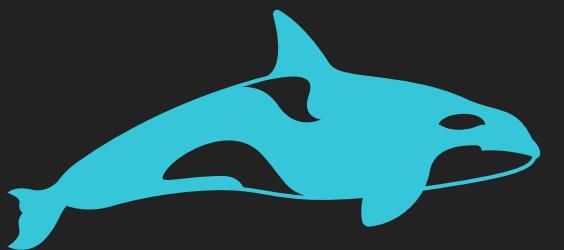
Rosco



Kayenta



Fiat



Orca



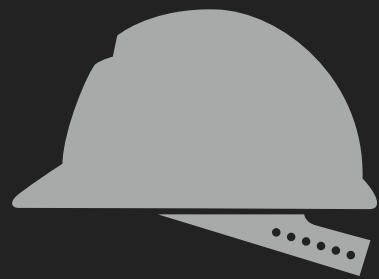
Halyard



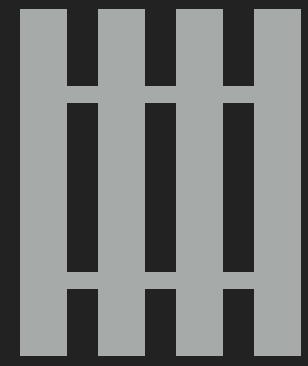
Clouddriver



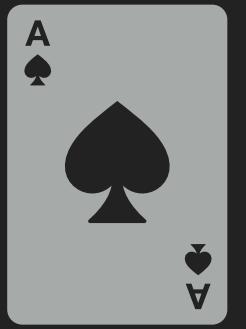
Echo



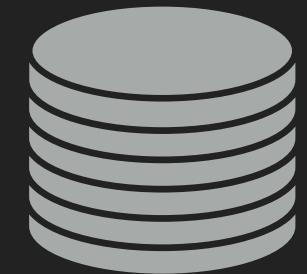
Igor



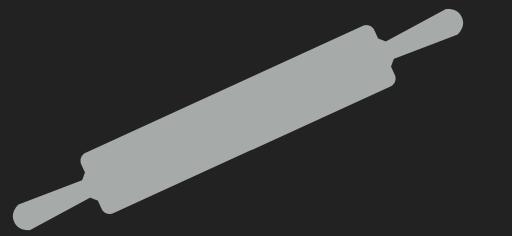
Gate



Deck



Front50



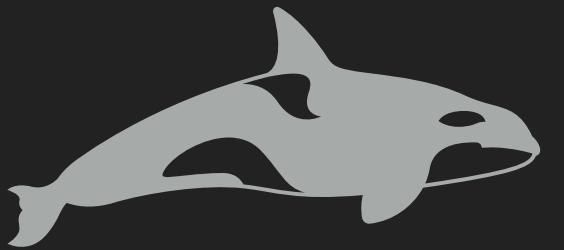
Rosco



Kayenta



Fiat

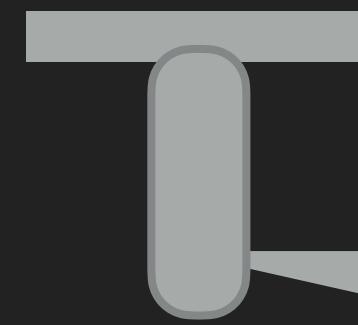


Orca

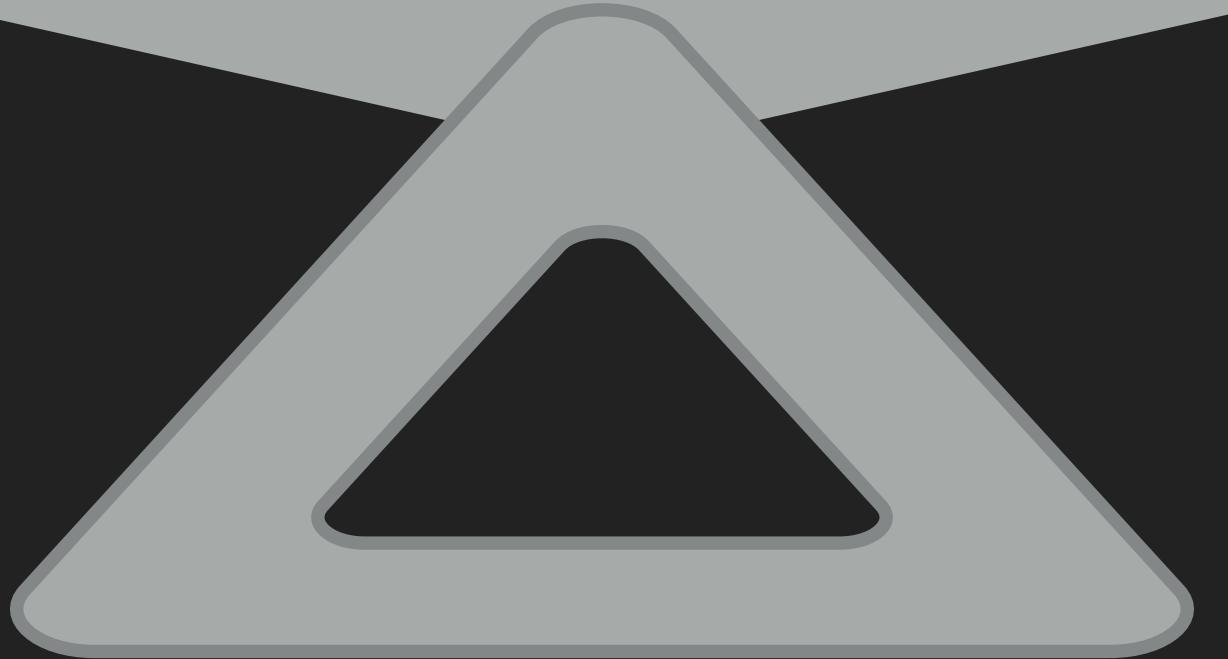
Kayenta is a platform for
Automated Canary Analysis (ACA)
and is used by Spinnaker to enable
automated canary deployments



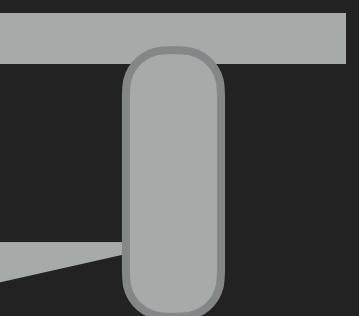
BUILD



COLLABORATE



BUY



Canary.



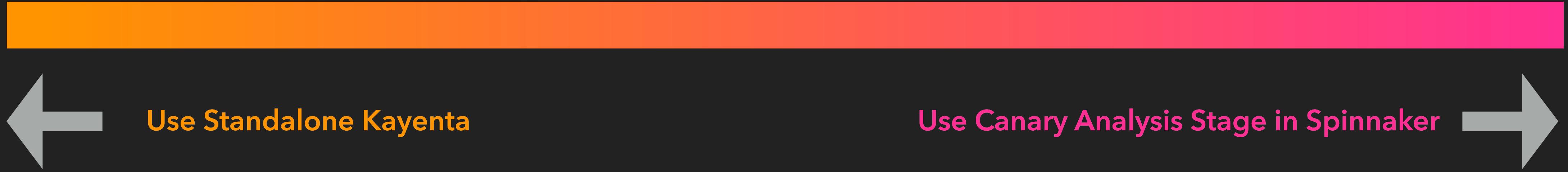
HOW TO CANARY

Use canary analysis
stage in Spinnaker

Use Standalone Kayenta

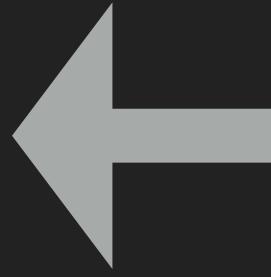


HOW SHOULD YOU CANARY?



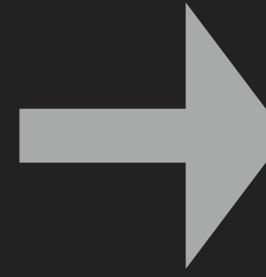
HOW SHOULD YOU CANARY?

Not using Spinnaker



Use Standalone Kayenta

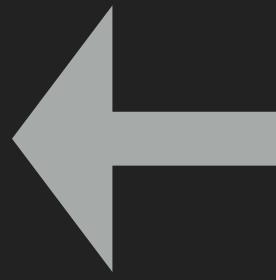
Use Canary Analysis Stage in Spinnaker



HOW SHOULD YOU CANARY?

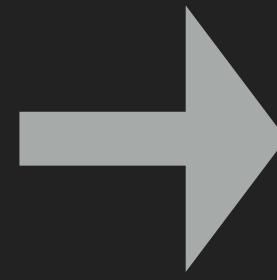
Not using Spinnaker

Use Spinnaker, but not every service will fully adopt it



Use Standalone Kayenta

Use Canary Analysis Stage in Spinnaker



HOW SHOULD YOU CANARY?

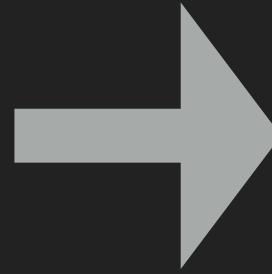
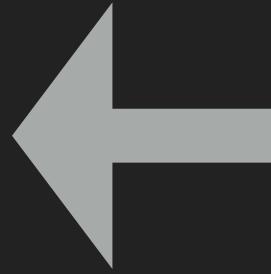
Not using Spinnaker

Use Spinnaker, but not every service will fully adopt it

Transitioning to Spinnaker

Use Standalone Kayenta

Use Canary Analysis Stage in Spinnaker



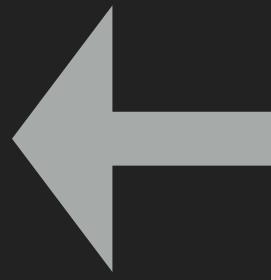
HOW SHOULD YOU CANARY?

Not using Spinnaker

Use Spinnaker, but not every service will fully adopt it

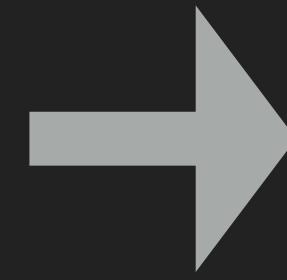
Transitioning to Spinnaker

Completely on Spinnaker



Use Standalone Kayenta

Use Canary Analysis Stage in Spinnaker



CANARYING WITH STANDALONE KAYENTA

1

Theory

How do canary deployments work with Standalone Kayenta?

2

Implementation

What can you use?
What do you have to build?

3

Experience

What does canarying look like for your engineers?



CANARYING WITH STANDALONE KAYENTA

1

Theory

How do canary deployments work with Standalone Kayenta?

2

Implementation

What can you use?
What do you have to build?

3

Experience

What does canarying look like for your engineers?



Define what makes an application healthy,
then measure and report it.



Define what makes an application healthy,
then measure and report it.

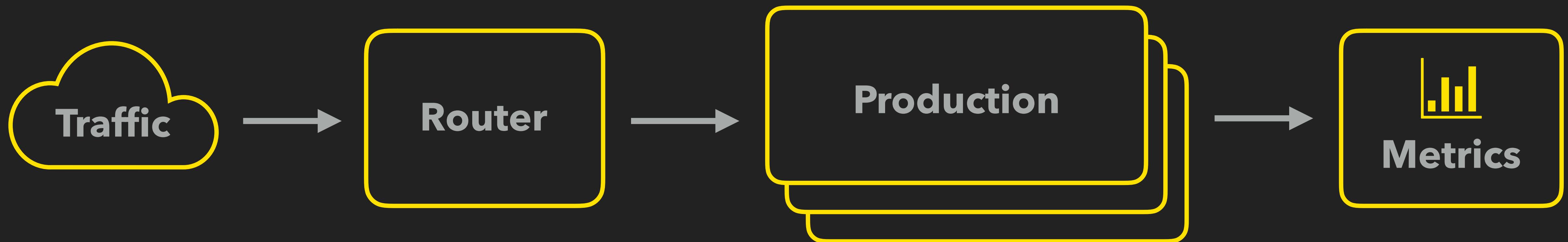
Saturation – How "full" your service is

Latency – The time it takes to service a request

Errors – The rate of requests that fail



PHASE 0 - STATE BEFORE CANARY DEPLOYMENT



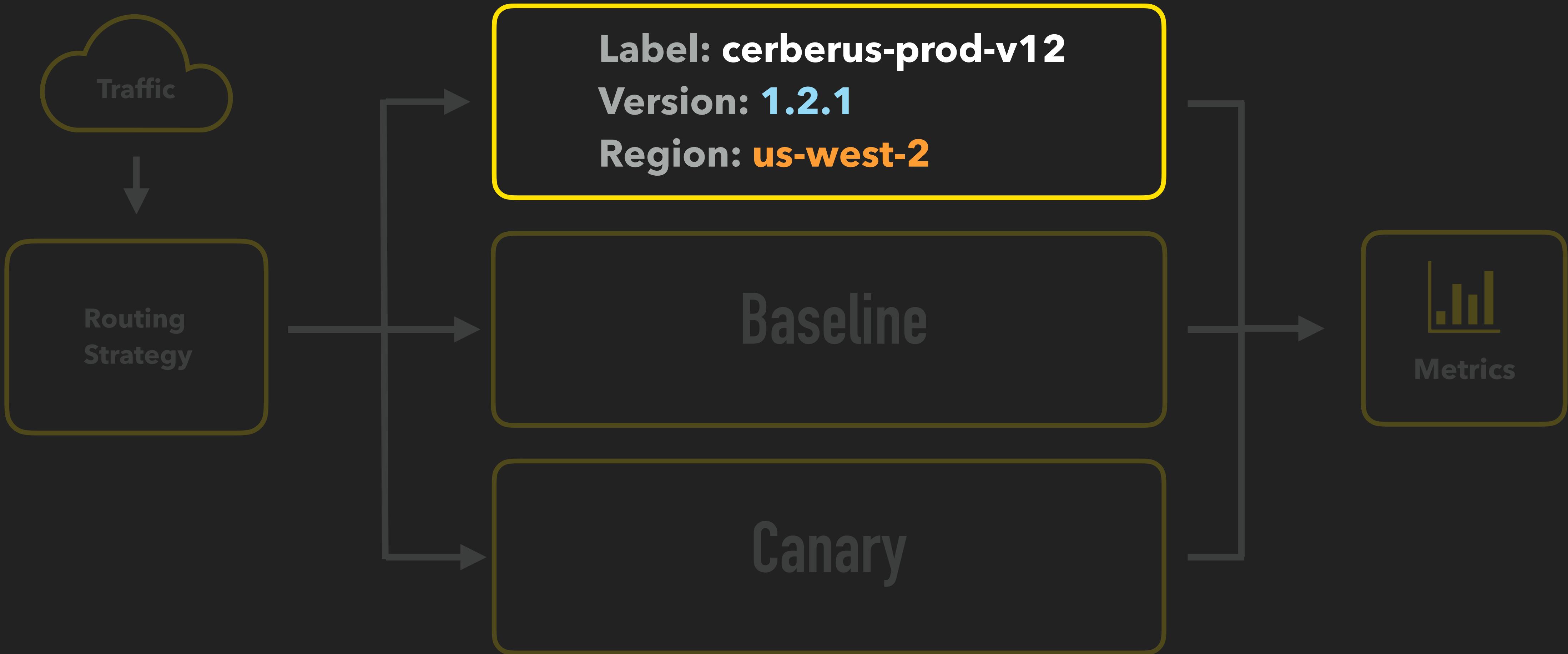
PHASE 1 - DEPLOYING THE BASELINE AND CANARY



PHASE 1 - DEPLOYING THE BASELINE AND CANARY



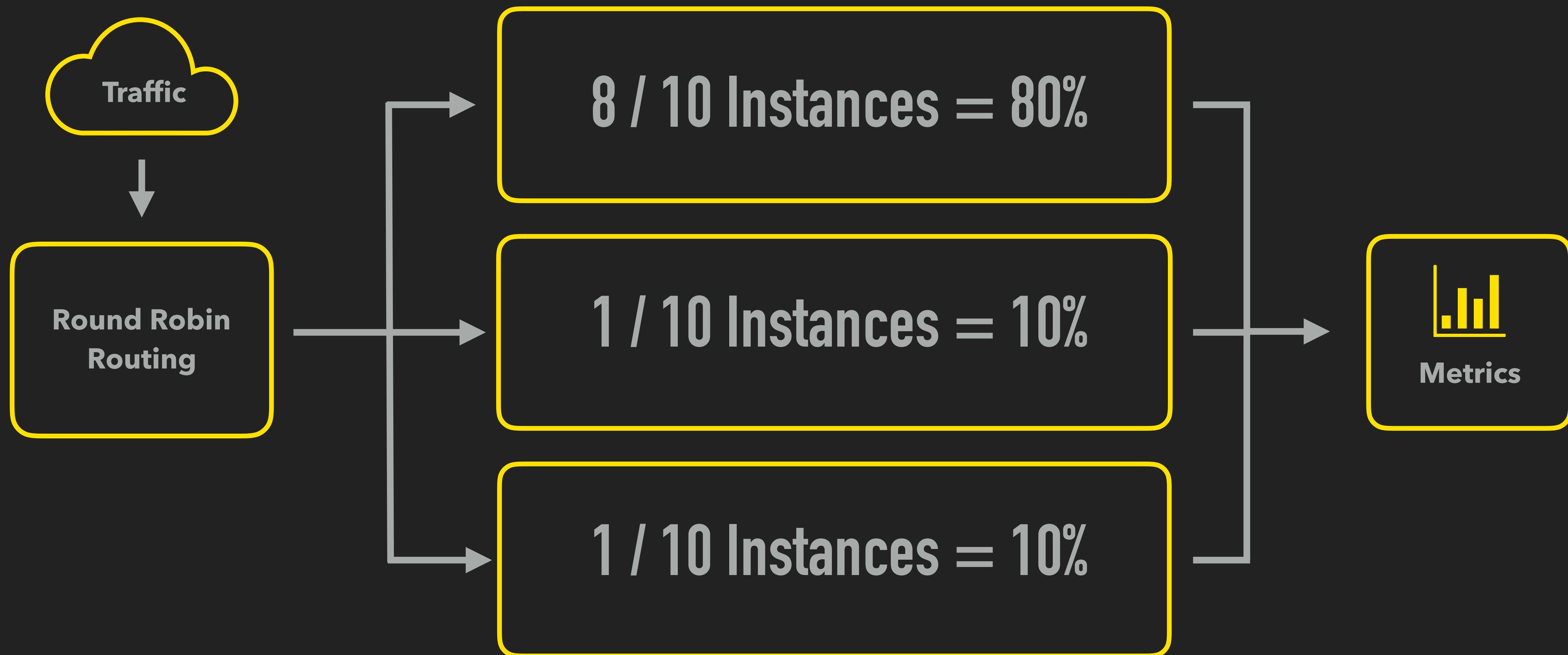
PHASE 1 - DEPLOYING THE BASELINE AND CANARY



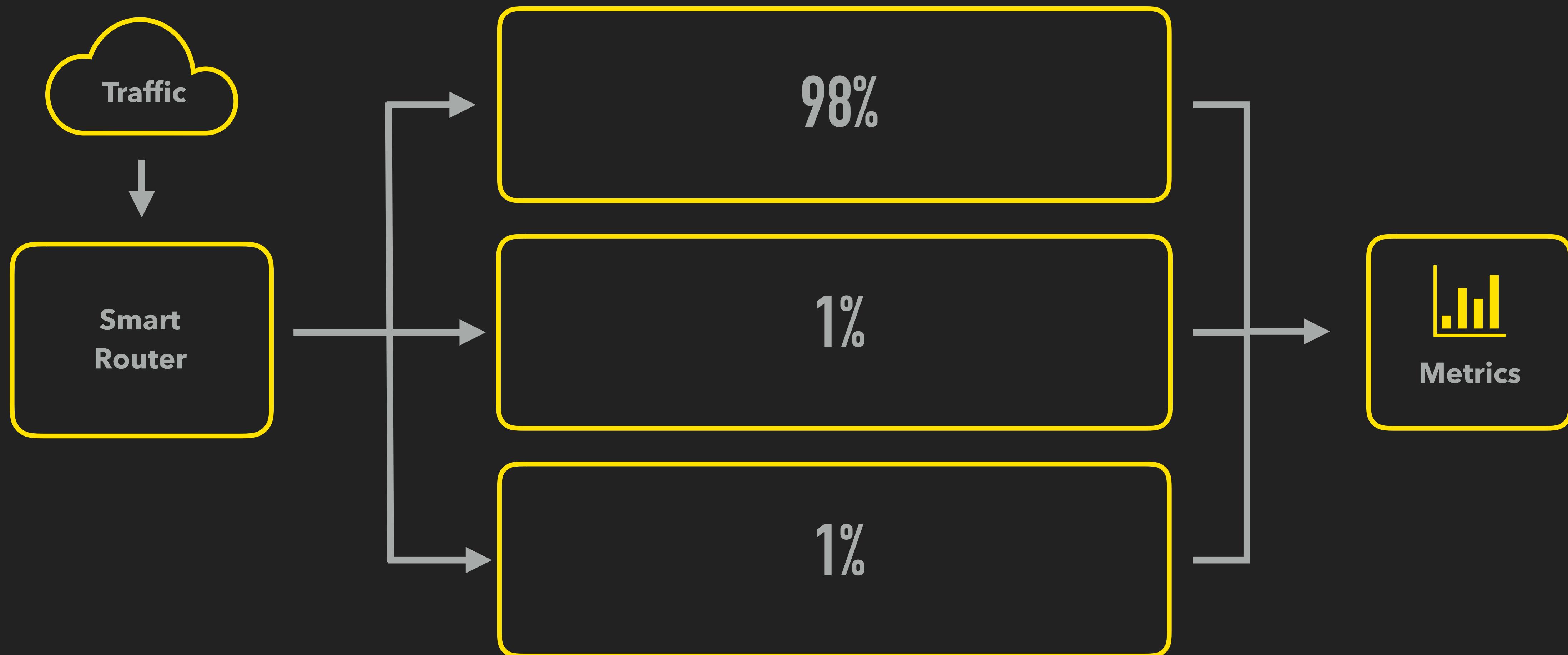
Baseline and Canary should receive equal amounts of traffic
so metrics such as counters can be compared equivalently



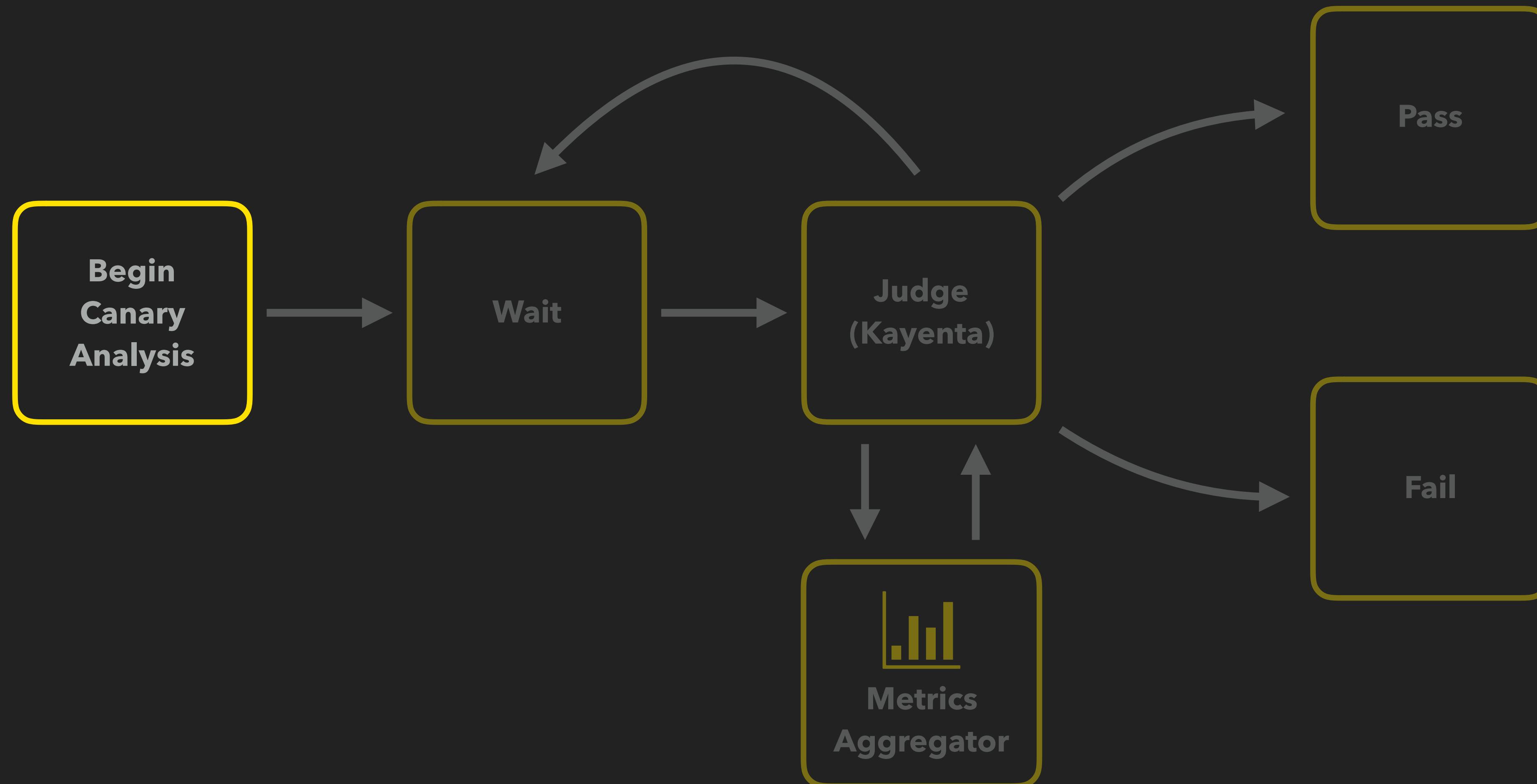
SIZE BASED ROUTING EXAMPLE (COURSE GRAIN)



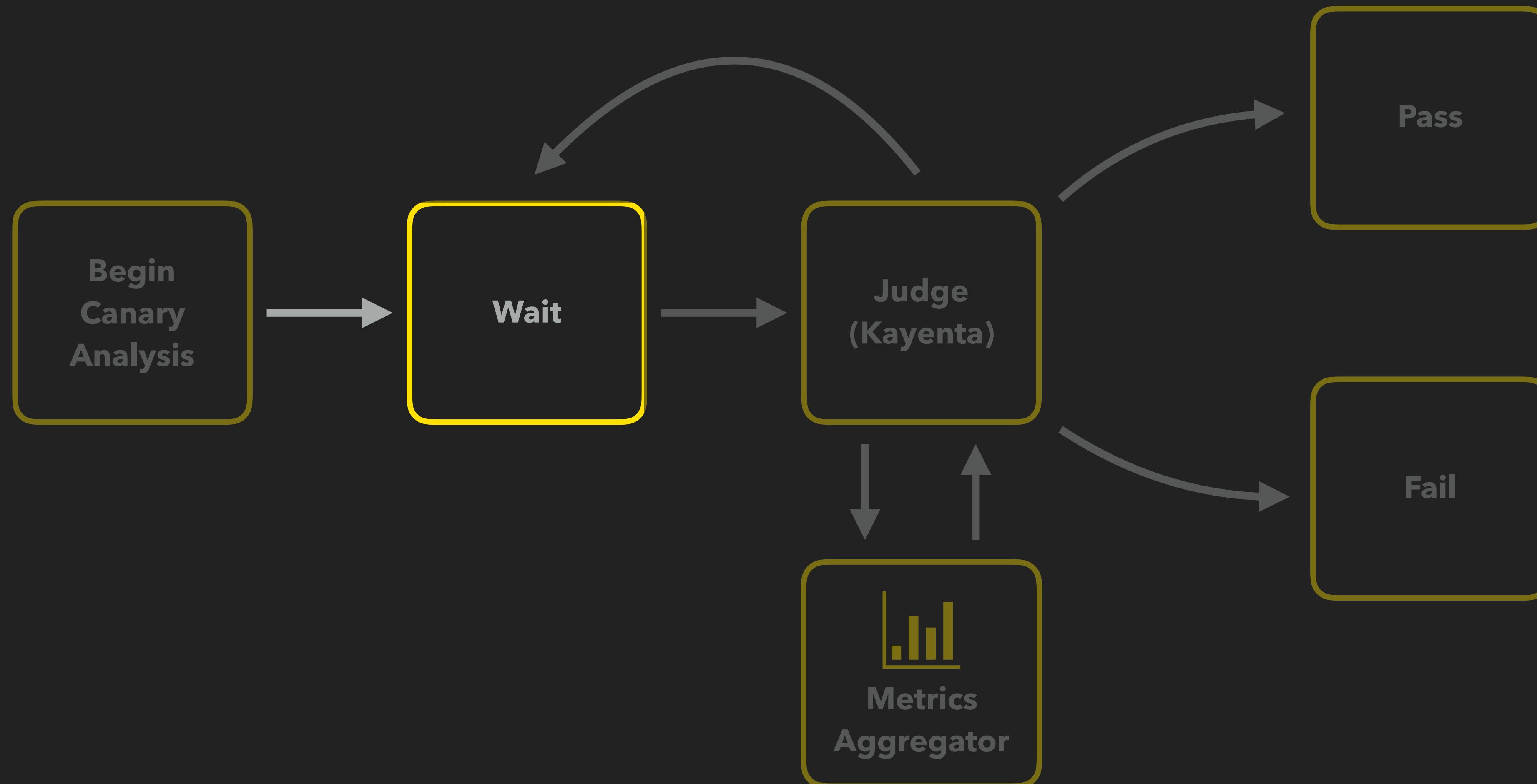
FINE GRAIN ROUTING EXAMPLE



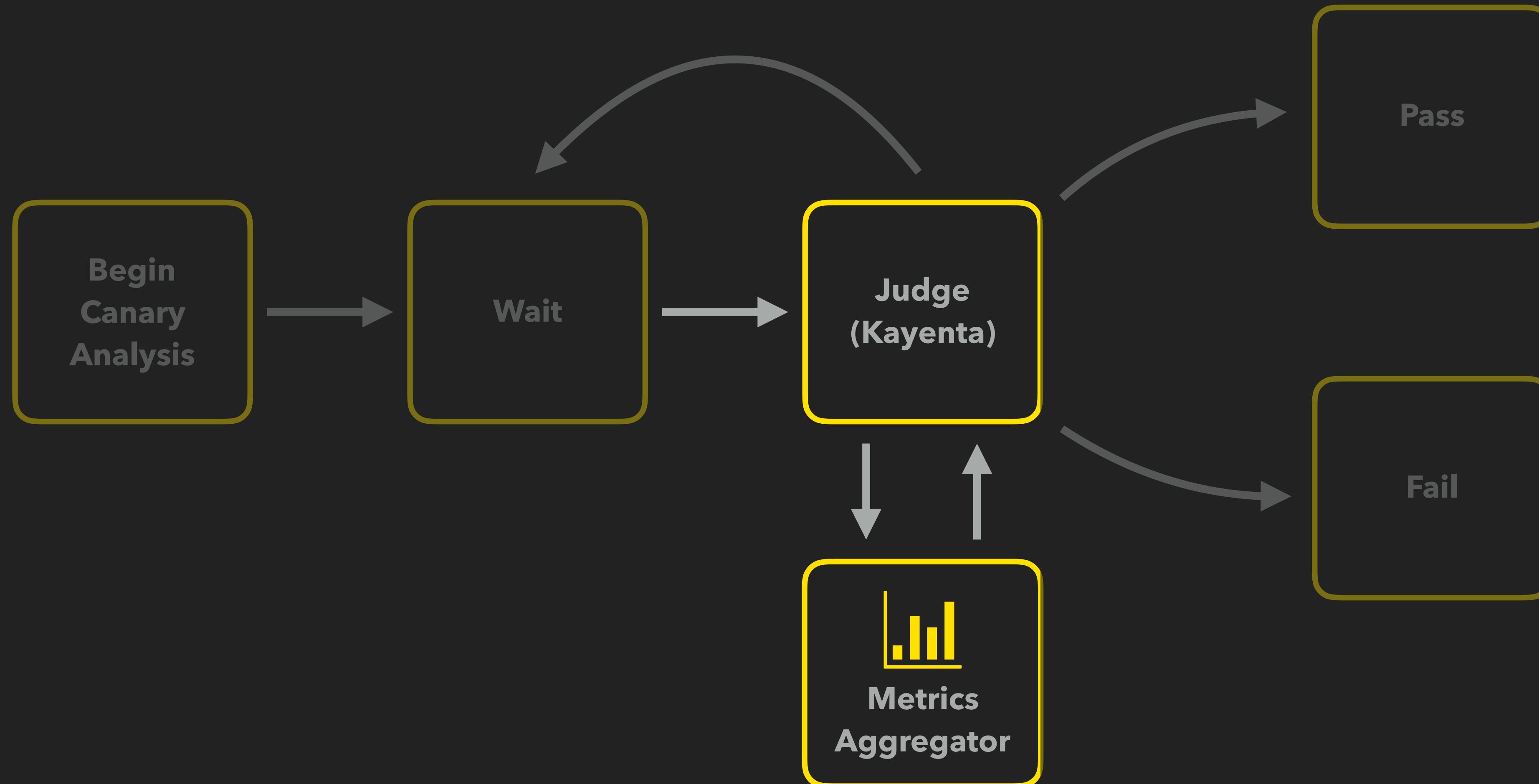
PHASE 2 - CANARY ANALYSIS



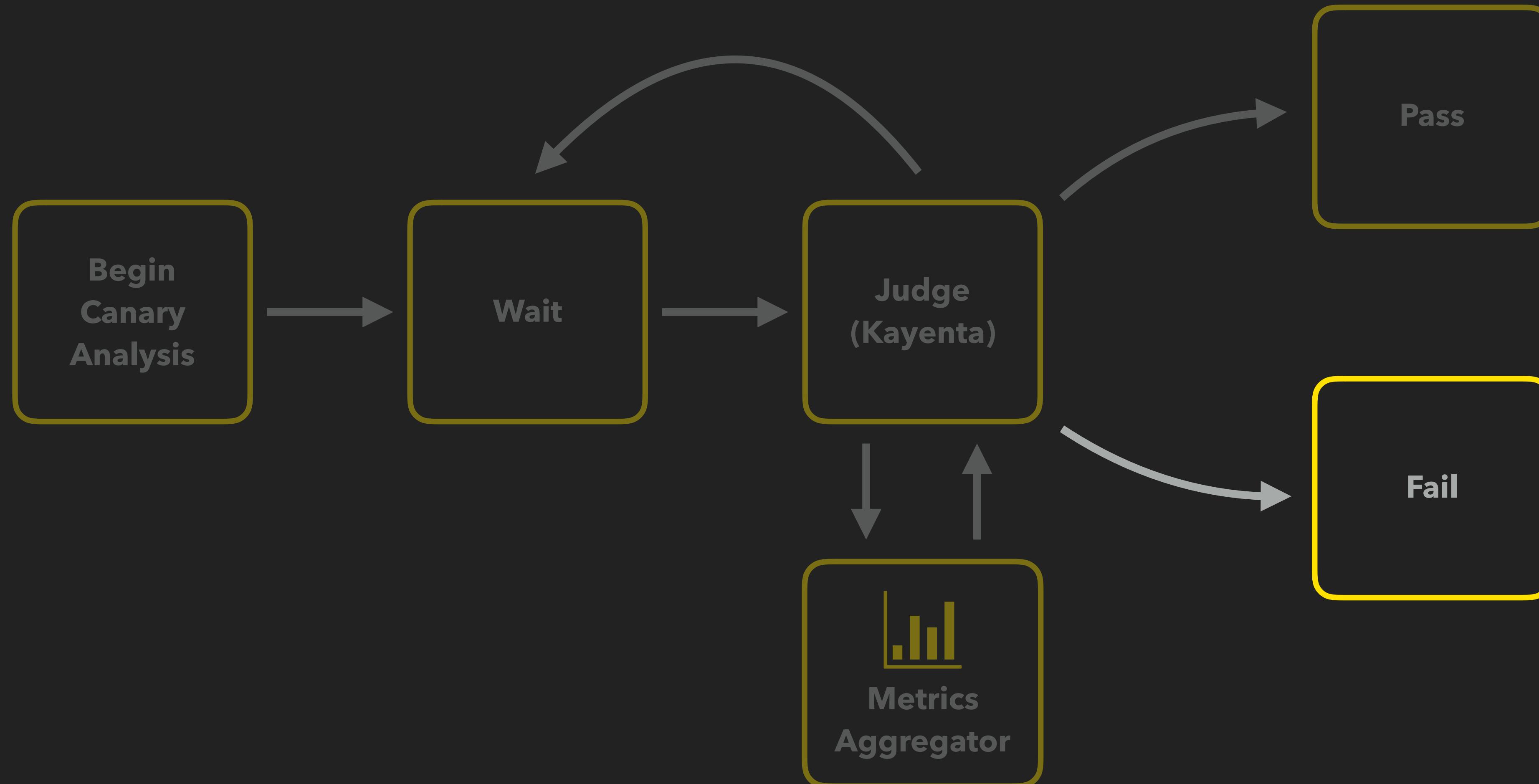
PHASE 2 - CANARY ANALYSIS



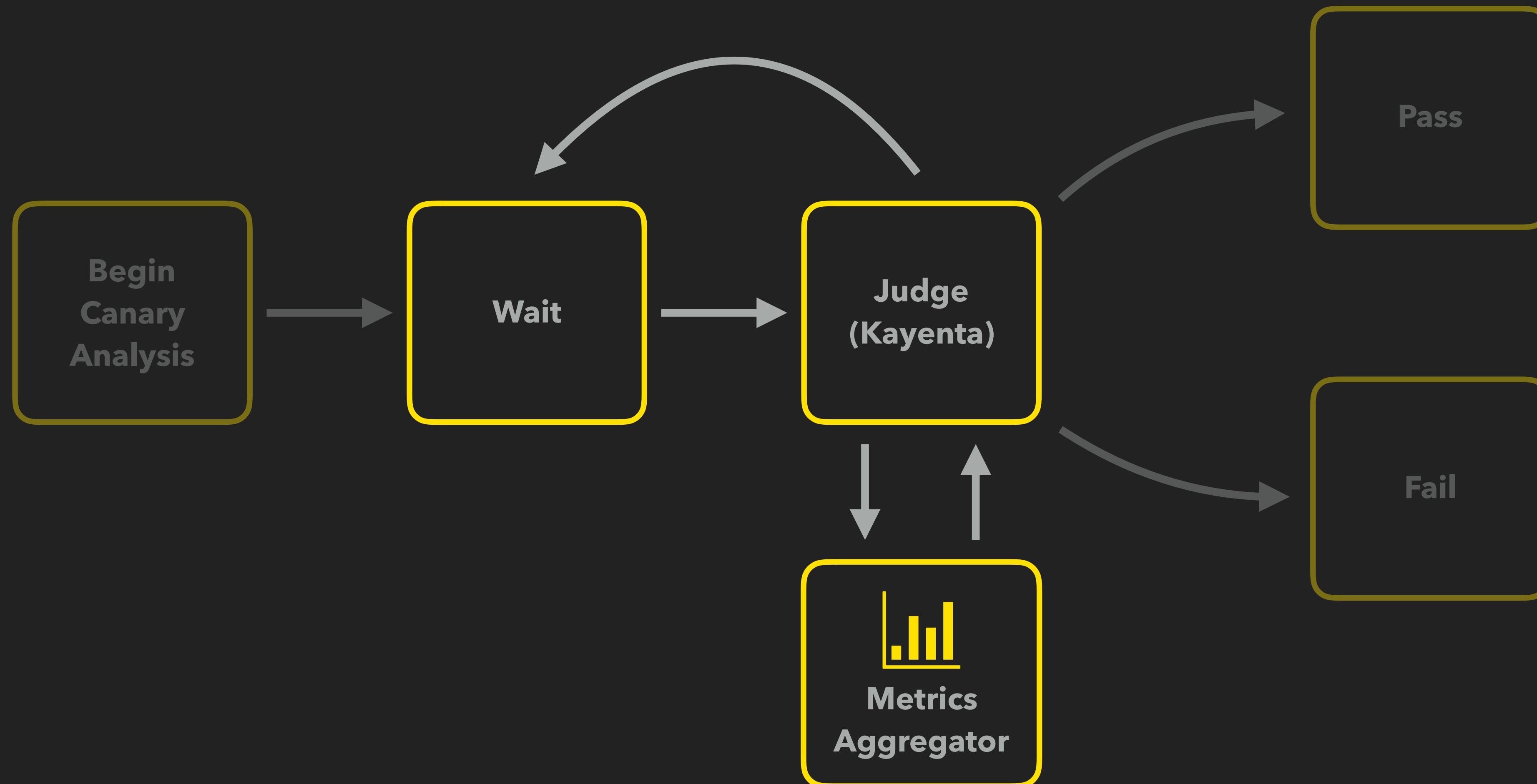
PHASE 2 - CANARY ANALYSIS



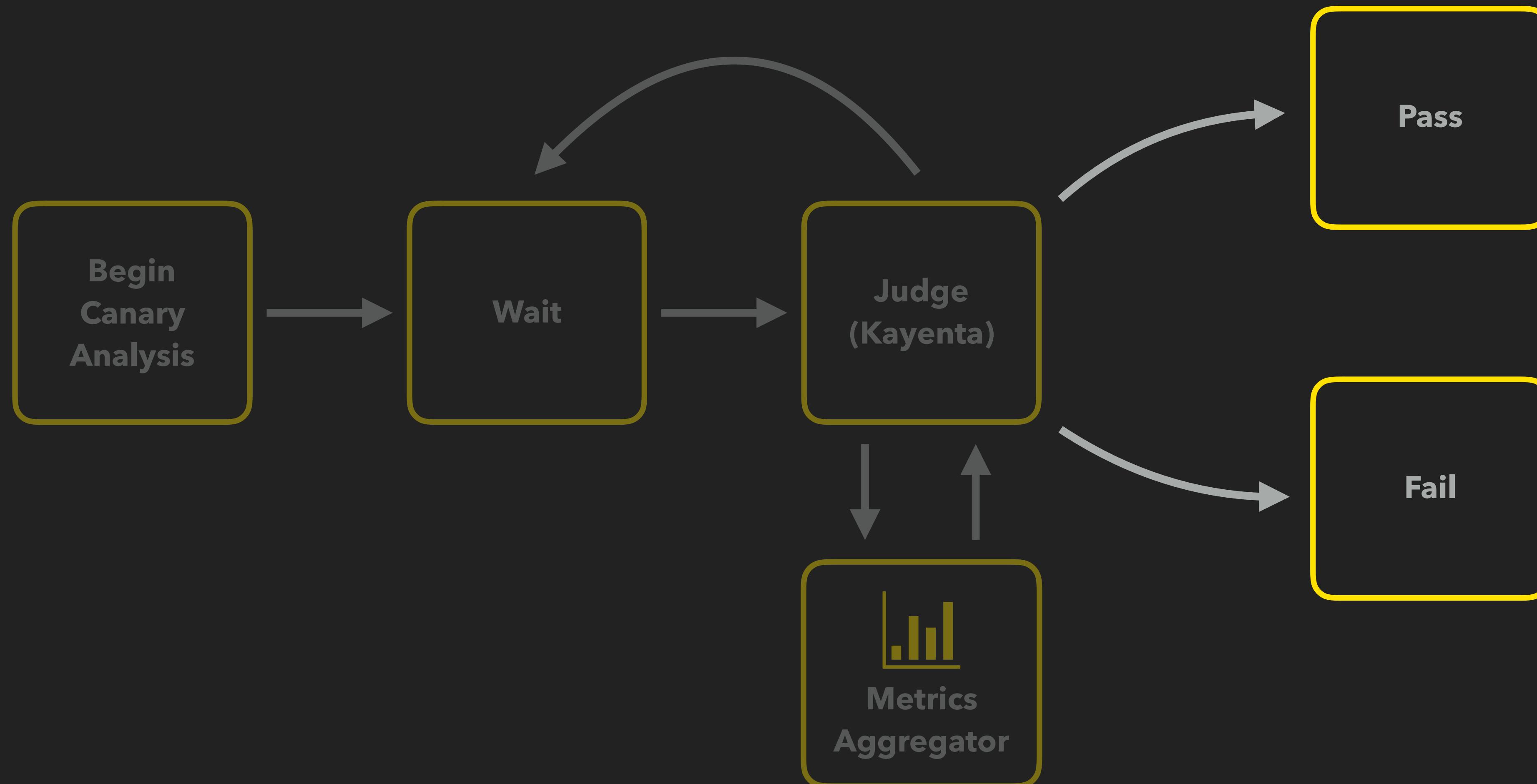
PHASE 2 - CANARY ANALYSIS



PHASE 2 - CANARY ANALYSIS



PHASE 2 - CANARY ANALYSIS



PROCESS RECAP

Report Metrics

Setup Infrastructure

Perform Canary Analysis



CANARYING WITH STANDALONE KAYENTA

1

Theory

How do canary deployments work with Standalone Kayenta?

2

Implementation

What can you use?
What do you have to build?

3

Experience

What does canarying look like for your engineers?





WHAT CAN YOU USE?





Referee

github.com/nike-inc/referee

Metadata

Metric Source Type

You must delete all metrics from the metrics section to change the metric source type

Configuration Name ?

Configuration Description ?

Metrics ?

[ALL](#) [ERRORS](#) [LATENCY](#) [SATURATION](#) [Add Group](#) ?

Metric Name	Groups	
Average Endpoint Latency	Latency	Edit Copy Delete
Bad Request Count	Errors	Edit Copy Delete
CPU Usage	Saturation	Edit Copy Delete
JVM Memory Heap used	Saturation	Edit Copy Delete
Server Error Count	Errors	Edit Copy Delete
Success Count	Errors	Edit Copy Delete

[Add Metric](#) ?

Scoring ?

Group Weights ?

Latency:

Errors: 70

Saturation: 30

Weight Total: 100 / 100

[Copy JSON to Clipboard](#)

[Save as File](#)

[Test Config](#)

Configuration Description

Cerberus Management Service Canary Configuration

Metrics

[ALL](#) [ERRORS](#) [LATENCY](#) [SATURATION](#) [Add Group](#) 

Metric Name	Groups	
Average Endpoint Latency	Latency	Edit Copy Delete
Bad Request Count	Errors	Edit Copy Delete
CPU Usage	Saturation	Edit Copy Delete
JVM Memory Heap used	Saturation	Edit Copy Delete
Server Error Count	Errors	Edit Copy Delete
Success Count	Errors	Edit Copy Delete



Metadata

Application Name

cerberus

Metrics Account

sfx-architecture

Storage Account

referee-s3-bucket

Thresholds

Marginal Threshold ⓘ

50

Pass Threshold ⓘ

75

Testing Type

Type

 A-A ⓘ A-B ⓘ

Scopes

Control

Scope ⓘ

cerberus-prod-baseline-v1

Location ⓘ

us-west-2

Step (s) ⓘ

60

Start Time (PST)

2019-10-16 16:10

Start Time (ISO)

2019-10-16T23:10:00.000Z

End Time (PST)

2019-10-16 19:10

End Time (ISO)

2019-10-17T02:10:00.000Z

Extended Scope Parameters

Add New Parameter ⓘ

Experiment

Scope ⓘ

cerberus-prod-canary-v1

Location ⓘ

us-west-2

Step (s) ⓘ

60

Start Time (PST)

2019-10-16 16:10

Start Time (ISO)

2019-10-16T23:10:00.000Z

End Time (PST)

2019-10-16 19:10

End Time (ISO)

2019-10-17T02:10:00.000Z

Extended Scope Parameters

Add New Parameter ⓘ

Testing Type

Type

A-A ?

A-B ?

Scopes

Control

Scope ?

cerberus-prod-baseline-v1

Location ?

us-west-2

Step (s) ?

60

Start Time (PST)

2019-10-16 16:10

Start Time (ISO)

2019-10-16T23:10:00.000Z

End Time (PST)

2019-10-16 19:10

End Time (ISO)

2019-10-17T02:10:00.000Z

Extended Scope
Parameters

Add New Parameter ?

Experiment

Scope ?

cerberus-prod-canary-v1

Location ?

us-west-2

Step (s) ?

60

Start Time (PST)

2019-10-16 16:10

Start Time (ISO)

2019-10-16T23:10:00.000Z

End Time (PST)

2019-10-16 19:10

End Time (ISO)

2019-10-17T02:10:00.000Z

Extended Scope
Parameters

Add New Parameter ?

Back to Canary Config

Run Manual Exec

Canary Report

Application

cms

User

anonymous

Metrics Account

sfx-architecture

Storage Account

referee-s3-bucket

Baseline

carebears-cms-master-cerberus-prod2-baseline-v001
us-west-2

Canary

carebears-cms-master-cerberus-prod2-experiment-v001
us-west-2

Time

9/27/2019, 1:32:53 PM
9/27/2019, 3:39:04 PM

PST ISO
PST ISO

Lifetime

126m

Interval

30m

Delay

5m

Step

10s

Go to Config

100
PASS100
PASS100
PASS100
PASS
 Fail (0) No Data (0) Pass (6)

Overview

Errors

100

Bad Request Count



Server Error Count



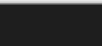
Success Count



Saturation

100

CPU Usage



JVM Memory Heap used



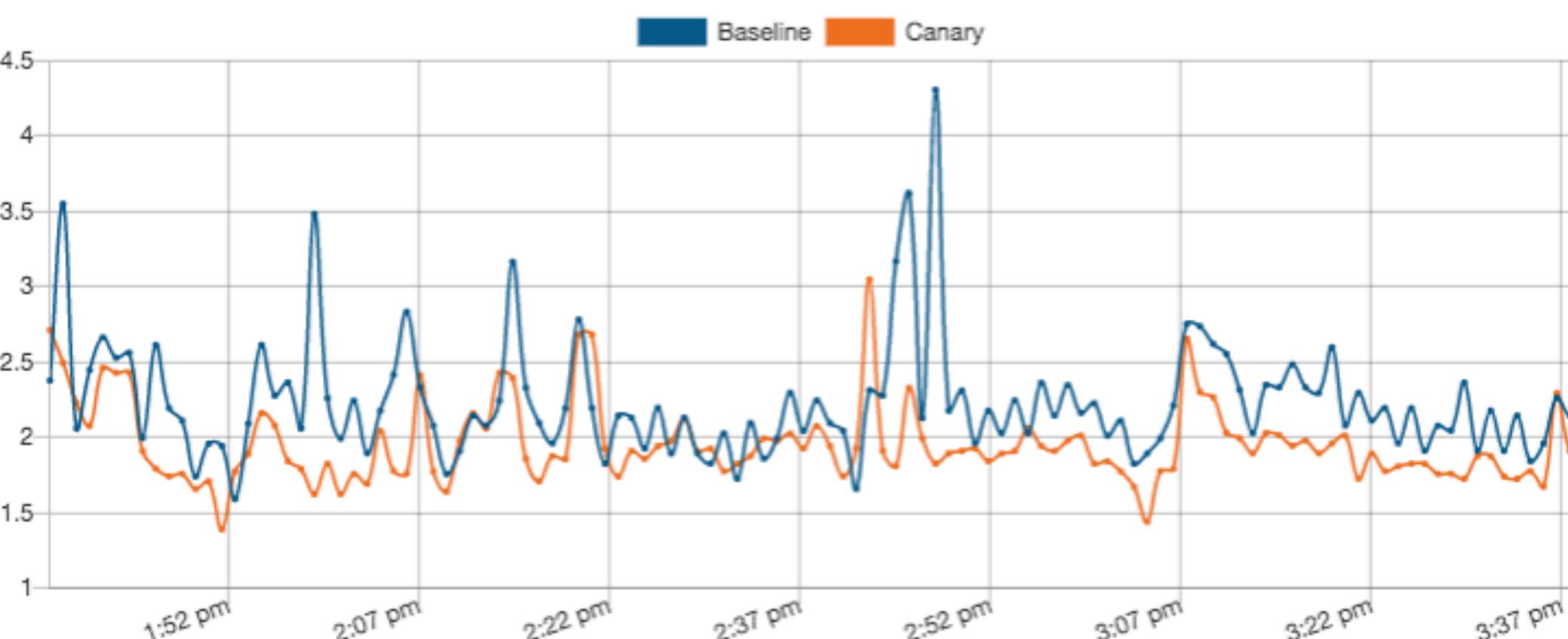
Latency

100

Average Endpoint Latency


● Fail
 ● No Data
 ● Pass
 ! Critical

CPU Usage



	Baseline	Canary
Count	119	120
Average	2.2237	1.9411
Min	1.5908	1.3903
Max	3.6187	2.7127
Std Dev	0.3514	0.2426

Baseline Query

```
data('cpu.utilization', filter=filter('server_group', 'carebears-cms-master-cerberus-prod2-baseline-v001') and
filter('server_region', 'us-west-2')).mean(by=['server_region', 'server_group']).publish()
```



[Home](#)[How Referee Works](#)[Canary Deployments](#)[Canarying Trade Offs](#)[Kayenta](#)[Guides](#)[Configuring Metrics for Referee](#)[Referee - EC2 B/G Pipeline Guide](#)[Referee - Custom Pipeline Guide](#)[Gatling Hive with Referee Guide](#)[Configuration](#)[Metrics Configuration](#)[Execution Configuration](#)[Templates](#)[Metrics Templates](#)[Execution Templates](#)[Best Practices](#)[Metrics](#)[Testing](#)

Canary Deployments

1. [Where does the term canary come from?](#)
2. [What are canary deployments?](#)
3. [Canary deployment flows](#)
4. [Why are canary deployments valuable?](#)
5. [When should engineers not use canary deployments?](#)

There are several types of [deployment strategies](#) for getting code out to production, such as big bang deployments, rolling deployments, and blue-green deployments. Canary deployments are another strategy to reduce the risk of introducing new software versions into production.

Where does the term canary come from?

- Canaries were brought into coal mines as a way to warn miners of danger
- If there was a dangerous gas leak, the canary would die first and the miners would know to evacuate
- Canaries are an early indicator of potential failure or danger

What are canary deployments?

- Canary deployments introduce a code change on a small percentage of your traffic
- The majority of your production traffic is protected from the change by continuing to run the current production code
- Metrics are collected on the behavior of both the baseline and the canary
- After a certain amount of time, the metrics are compared to evaluate the canary
- If the canary passes, the code change rolls out to all of production
- If the canary fails, the changes roll back

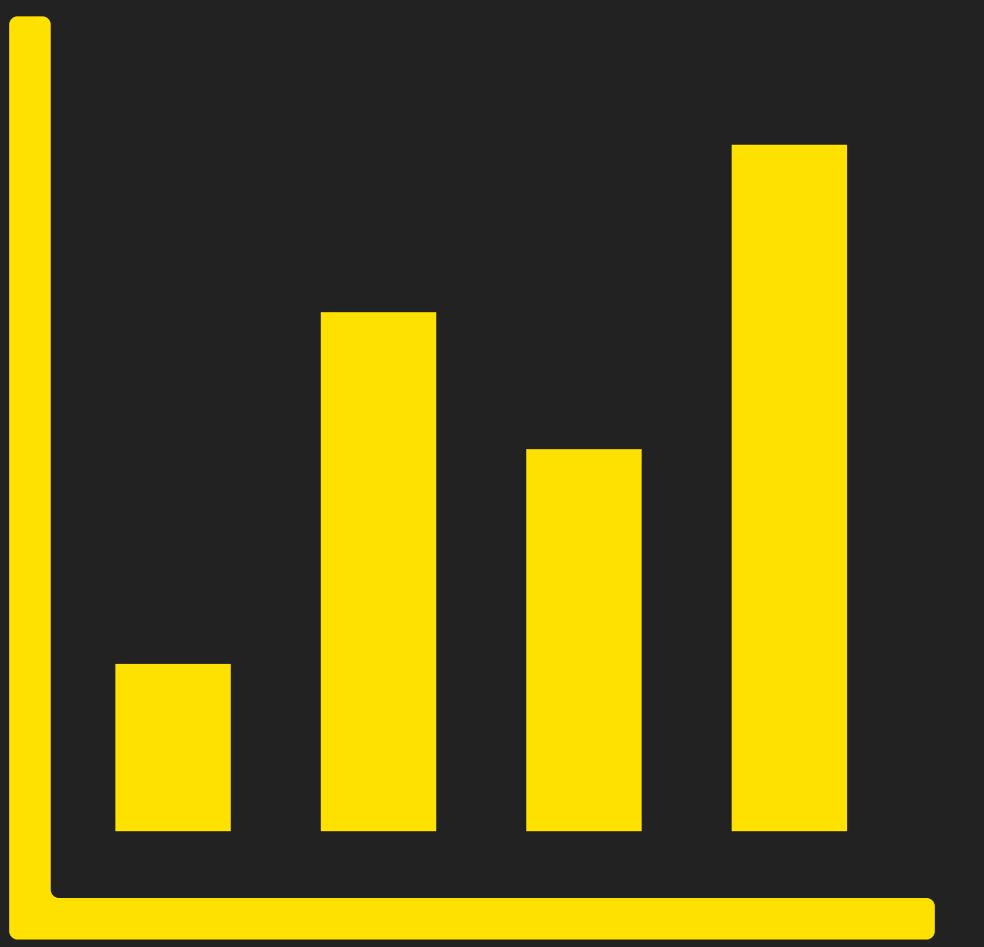
Canary deployment flows

Production and Canary

Here is an example of a simple canary deployment flow where traffic is routed between production and canary.



OPEN SOURCE CONTRIBUTIONS



SignalFx Integration



OPEN SOURCE CONTRIBUTIONS



New Relic
Integration Support

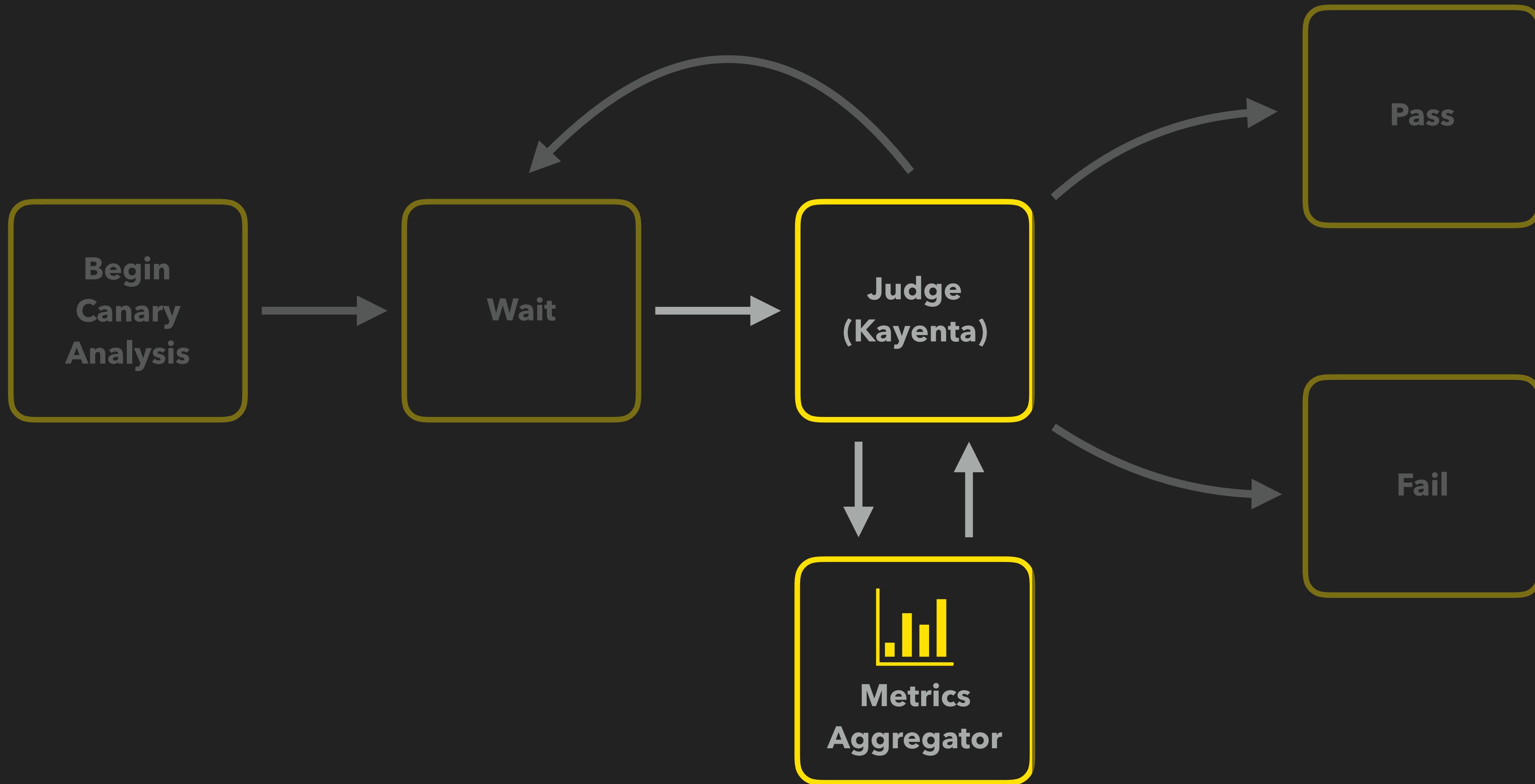


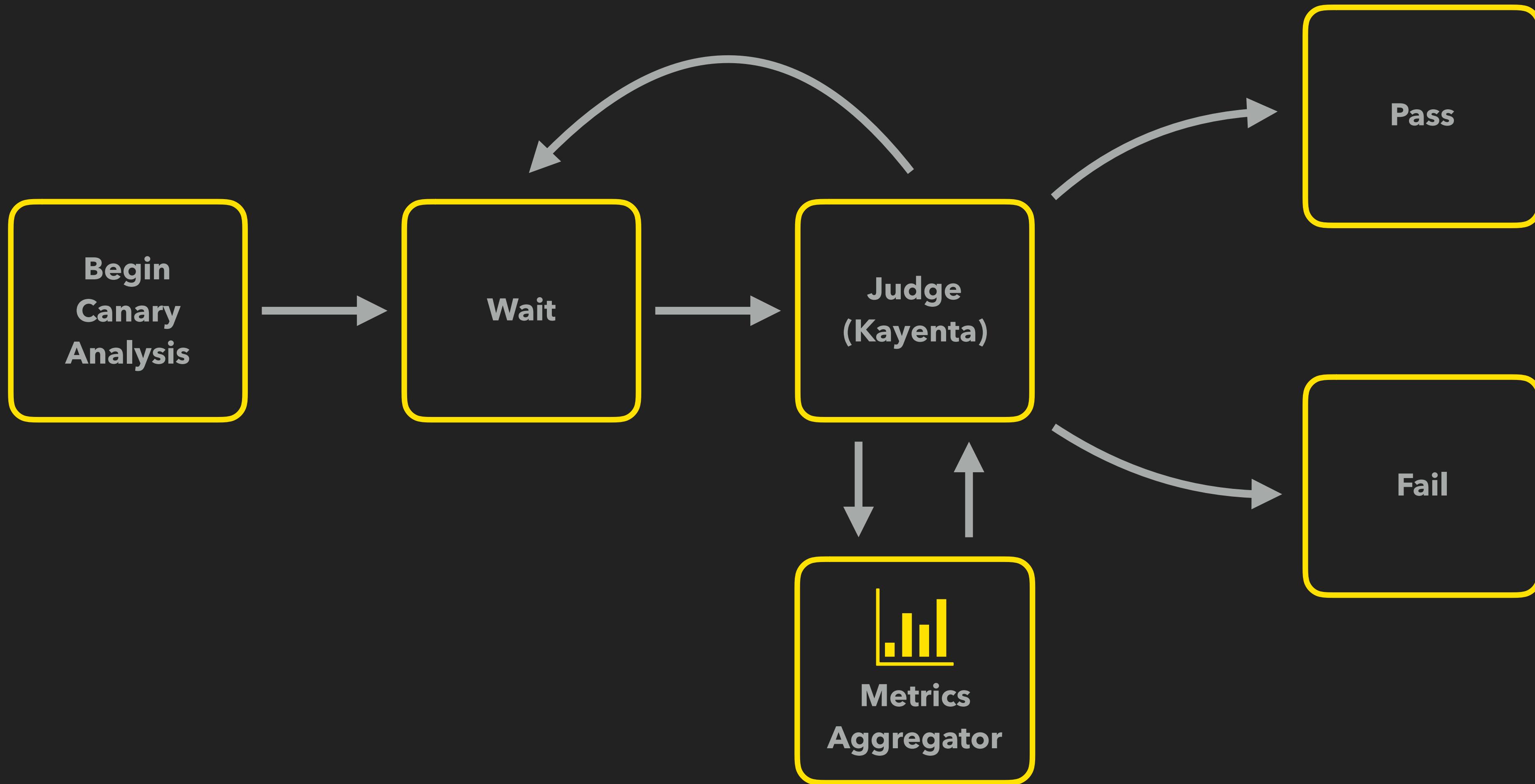
OPEN SOURCE CONTRIBUTIONS



Kayenta Standalone
endpoint in API



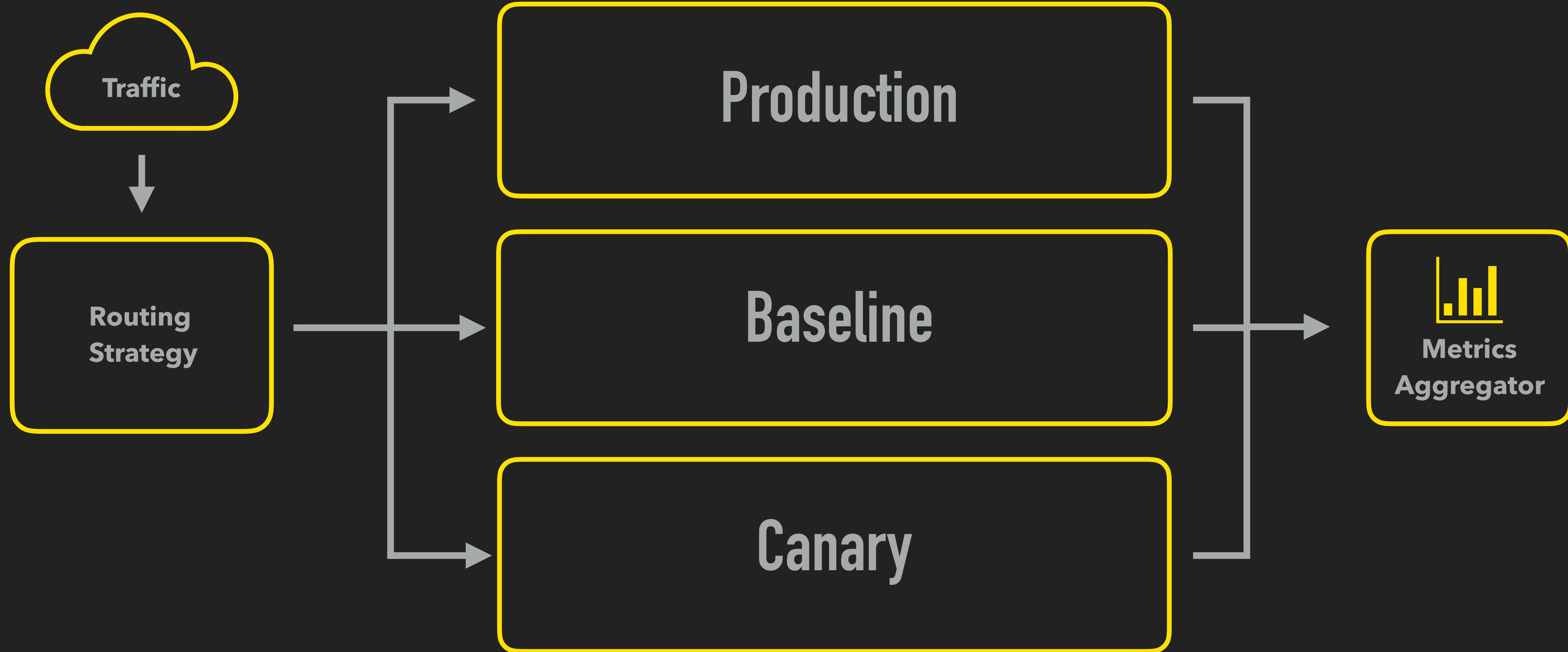




A black and white photograph of four female runners jogging on a city street at night. They are wearing athletic gear, including tank tops and shorts. The scene is set against a backdrop of city lights, buildings, and a train in the distance.

WHAT DO YOU HAVE TO BUILD?

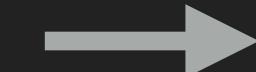




Server Group: cerberus-prod-v12
Version: 1.2.1
Region: us-west-2
Instance Count: 8



```
{  
  "name": "jmx_memory.used",  
  "value": 1400,  
  "dimensions": {  
    "server_group": "cerberus-prod-v12",  
    "region": "us-west-2",  
    "version": "1.2.1"  
  }  
}
```




Metrics
Aggregator





Referee Bot APP 2:55 PM

Automated Platform for Canary Analysis

Application

cms

Result

Fail

Message

Final canary score 85.0 is below the pass score threshold.

[View Report](#)

[Go to Build](#)



Referee Bot APP 3:39 PM

Automated Platform for Canary Analysis

Application

cms

Result

Pass

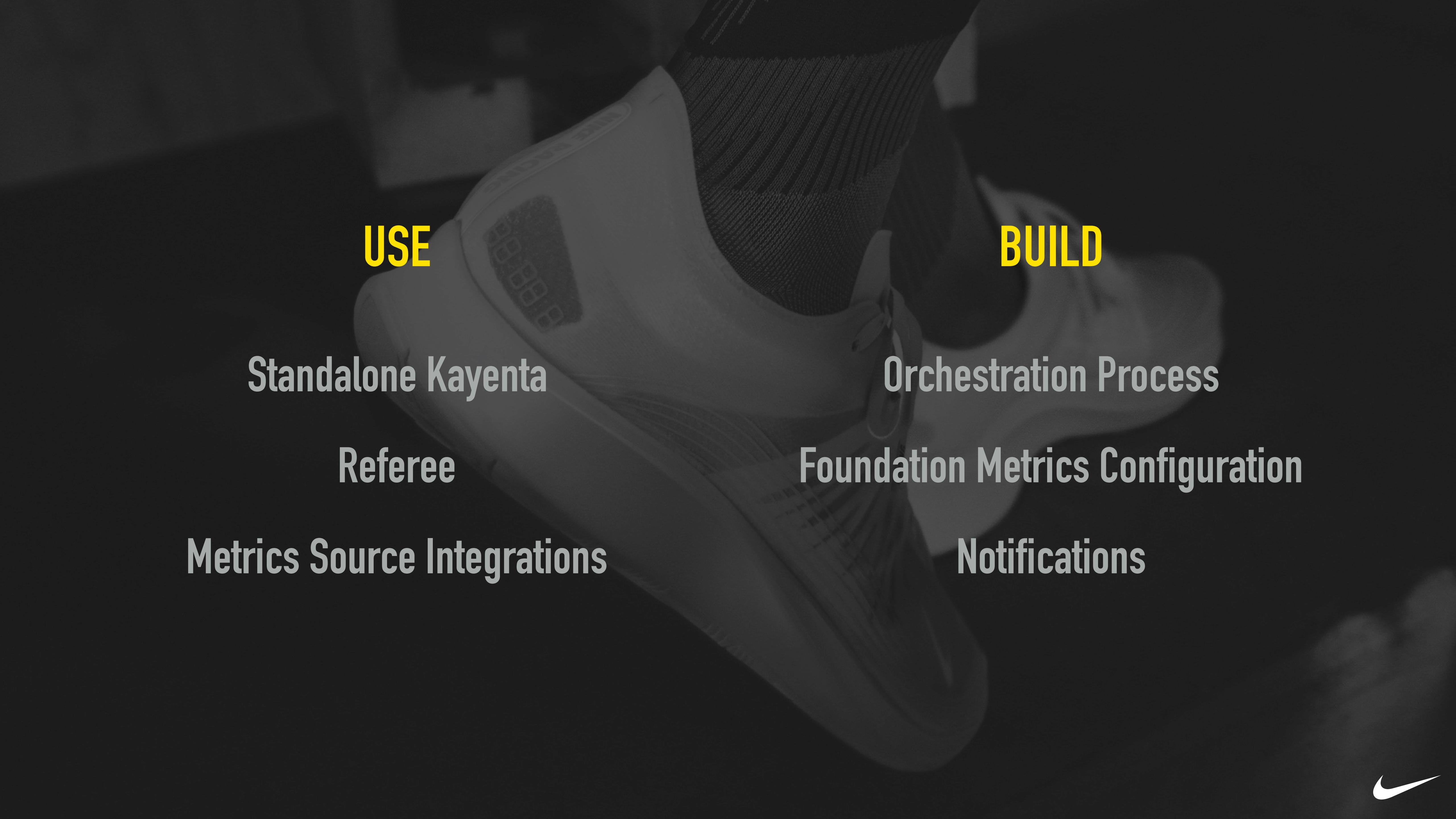
Message

Final canary score 100.0 met or exceeded the pass score threshold.

[View Report](#)

[Go to Build](#)





USE

Standalone Kayenta

Referee

Metrics Source Integrations

BUILD

Orchestration Process

Foundation Metrics Configuration

Notifications



CANARYING WITH STANDALONE KAYENTA

1

Theory

How do canary deployments work with Standalone Kayenta?

2

Implementation

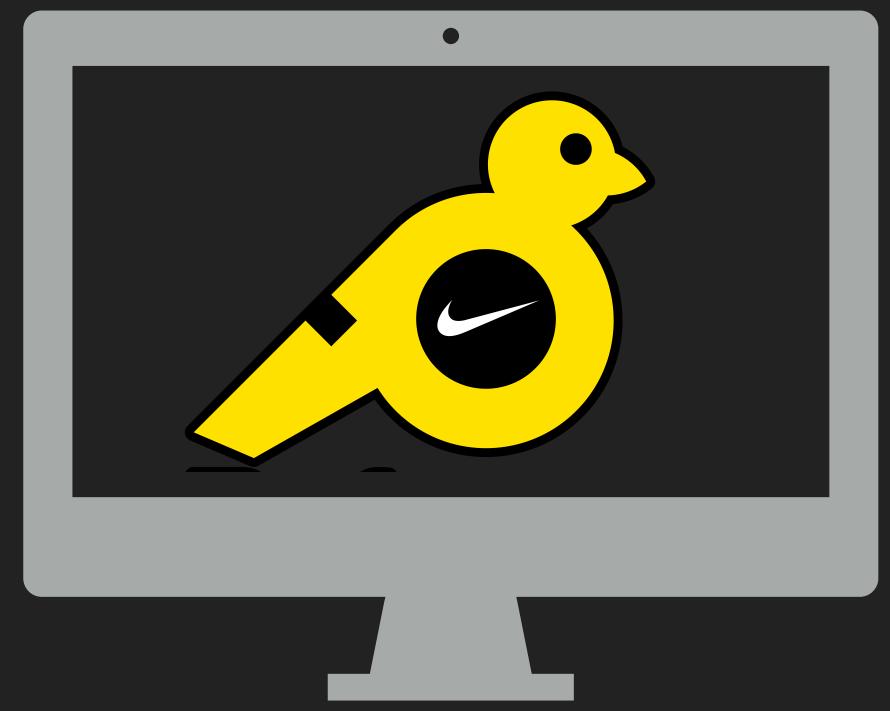
What can you use?
What do you have to build?

3

Experience

What does canarying look like for your engineers?





nike-inc.github.io/cerberus



CANARYING EXPERIENCE

Instrument metrics

Pipeline integration

Create canary config

Test config

Run canary deployments





Referee Bot APP 2:55 PM

Automated Platform for Canary Analysis

Application

cms

Environment

prod-west

Result

Fail

Message

Final canary score 85.0 is below the pass score threshold.

[View Report](#)

[Go to Build](#)







Referee Bot APP 2:55 PM

Automated Platform for Canary Analysis

Application

cms

Environment

prod-west

Result

Fail

Message

Final canary score 78.0 is below the pass score threshold.

[View Report](#)

[Go to Build](#)



Canary Report

Application

cms

User

anonymous

Metrics Account

sfx-architecture

Storage Account

referee-s3-bucket

Baseline

carebears-cms-master-cerberus-prod2-baseline-v1

us-west-2

Canary

carebears-cms-master-cerberus-prod2-experiment-v1

us-west-2

Time

2019-08-07T22:09:57.503Z

PDT ISO

2019-08-07T22:57:56.553Z

PDT ISO

Lifetime

48m

Interval

15m

Delay

1m

Step

10s

Go to Config

89
MARGINAL78
MARGINAL78
FAIL Fail (2) No Data (0) Pass (7)

Overview

Errors

100

Bad Request Count



Server Error Count



Success Count



Saturation

33

CPU Usage



JVM Memory Heap used



Memory Usage



Latency

100

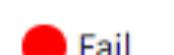
Average Endpoint Latency



KMS Auth Latency



Secure Data Read Latency



Fail



No Data



Pass



! Critical

Run

3 of 3

8/7/2019, 3:10:57 PM

PDT ISO

8/7/2019, 3:55:57 PM

PDT ISO



2

Failures

0

No Data

7

Passes



Canary Report

Application

cms

User

anonymous

Metrics Account

sfx-architecture

Storage Account

referee-s3-bucket

Baseline

carebears-cms-master-cerberus-prod2-baseline-v1

us-west-2

Canary

carebears-cms-master-cerberus-prod2-experiment-v1

us-west-2

Time

2019-08-07T22:09:57.503Z

PDT ISO

2019-08-07T22:57:56.553Z

PDT ISO

Lifetime

48m

Interval

15m

Delay

1m

Step

10s

Go to Config

89

MARGINAL

78

MARGINAL

78

FAIL

 Fail (2) No Data (0) Pass (7)

Overview

Errors

100

Bad Request Count



Server Error Count



Success Count



Saturation

33

CPU Usage



JVM Memory Heap used



Memory Usage



Latency

100

Average Endpoint Latency



KMS Auth Latency

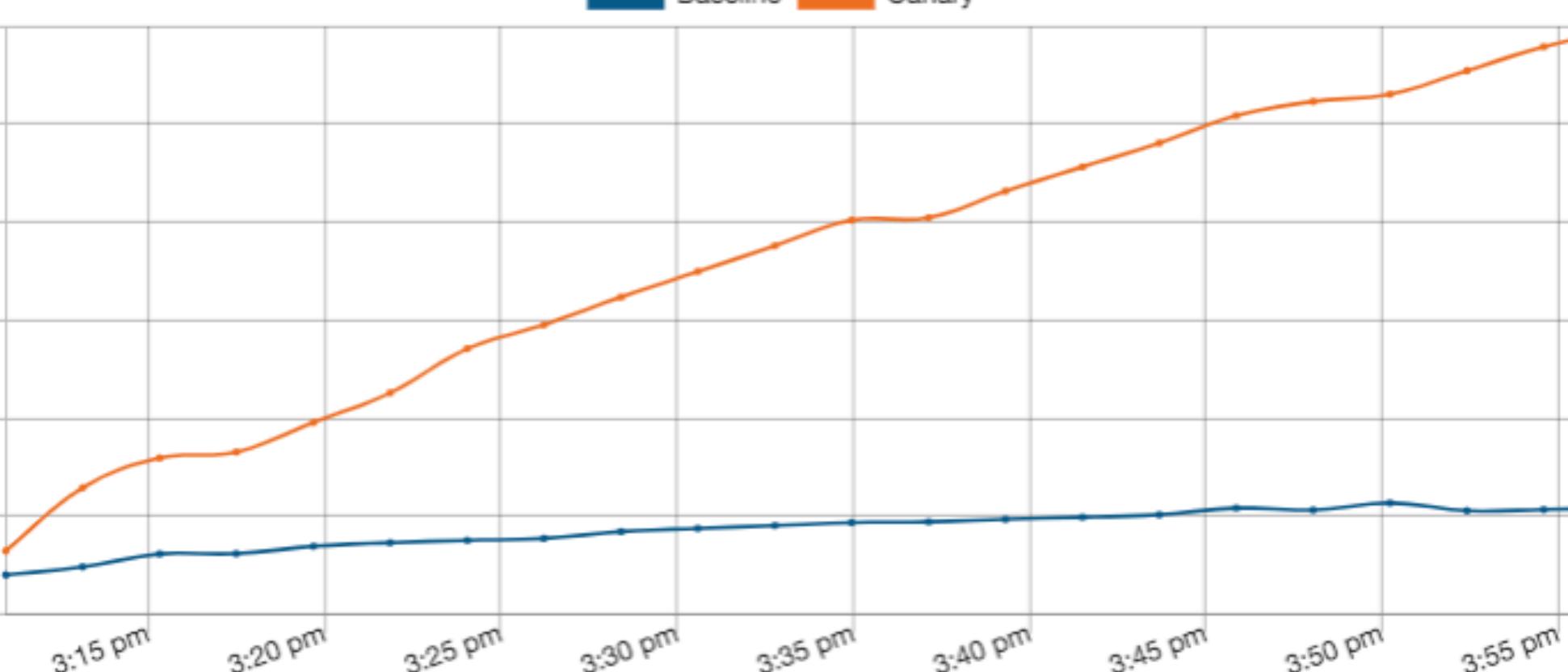


Secure Data Read Latency



Memory Usage

Baseline Canary



	Baseline	Canary
Count	22	22
Average	34.3381	48.2396
Min	32.0196	33.2442
Max	35.6853	59.8242
Std Dev	1.035	7.9331

Baseline Query



Canary Report

Application

cms

User

anonymous

Metrics Account

sfx-architecture

Storage Account

referee-s3-bucket

Baseline

- carebears-cms-master-cerberus-test2-baseline-v1
- us-west-2

Canary

- carebears-cms-master-cerberus-test2-experiment-v1
- us-west-2

Time

- 8/7/2019, 11:36:02 PM
- PST ISO
- 8/8/2019, 2:42:10 AM
- PST ISO

Lifetime

186m

Interval

30m

Delay

5m

Step

10s

[Go to Config](#)100
PASS100
PASS100
PASS100
PASS100
PASS100
PASS Fail (0) No Data (0) Pass (9)

Overview

Errors

100

Bad Request Count

!

Server Error Count

!

Success Count

!

Saturation

100

CPU Usage

!

JVM Memory Heap used

!

Memory Usage

!

Latency

100

Average Endpoint Latency

!

KMS Auth Latency

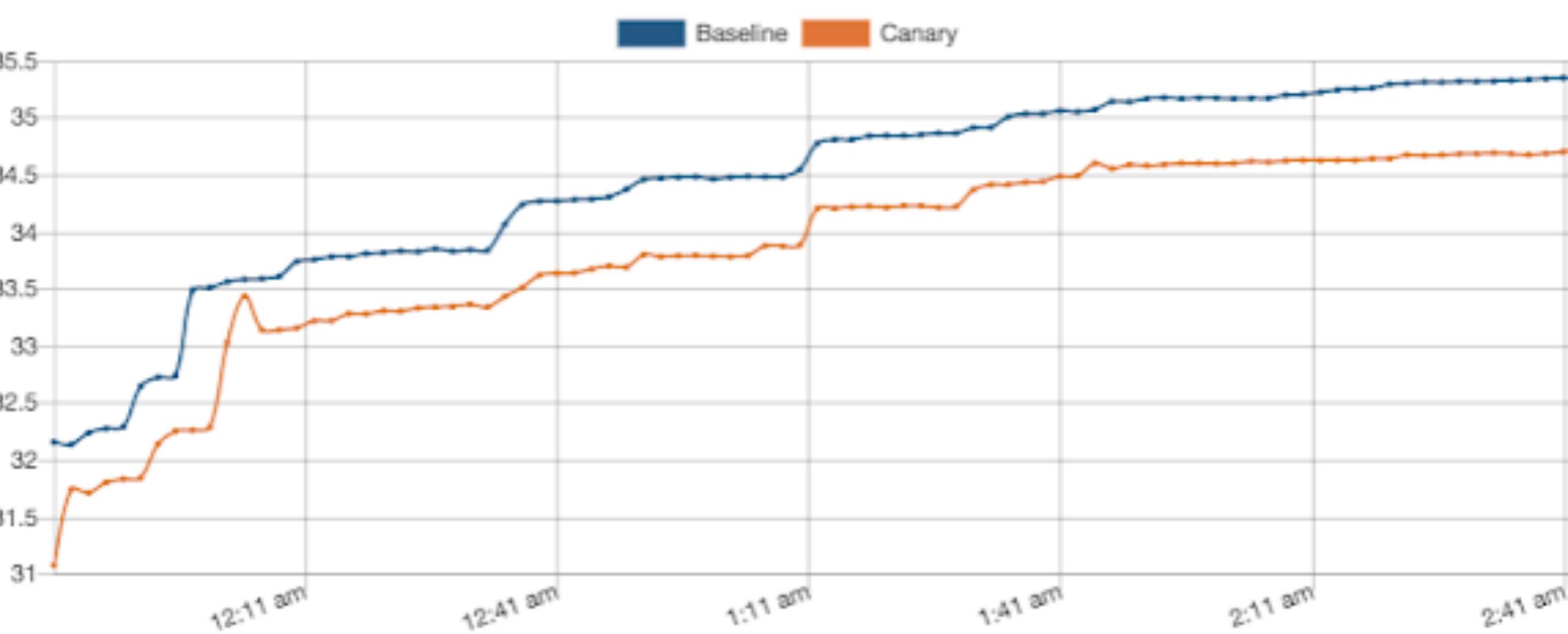
!

Secure Data Read Latency

!

● Fail● No Data● Pass! Critical

Memory Usage



	Baseline	Canary
Count	90	90
Average	34.4629	33.8572
Min	32.1424	31.0848
Max	35.3584	34.7475
Std Dev	0.8664	0.875



Canarying prevented a potential
production outage



Canary.



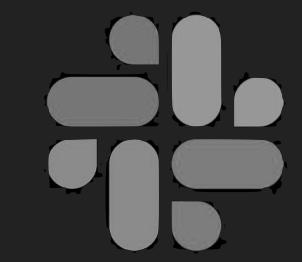


WHAT CAN YOU DO TODAY?





Check out the code:
github.com/spinnaker/kayenta
github.com/nike-inc/referee



Join us on Spinnaker Slack #kayenta:
join.spinnaker.io



Justin.Field@nike.com Melana.Hammel@nike.com



jobs.nike.com



RESOURCES

- ▶ [Canary analysis: Lessons learned and best practices from Google and Waze](#)
- ▶ [Best practices for configuring canary](#)
- ▶ [9 Reliability-Based Best Practices for Canary Deploys](#)
- ▶ [Canarying Well: Lessons Learned from Canarying Large Populations](#)
- ▶ [Matt Duffler & Michael Graff: Performing Automated Canary Analysis with Kayenta and Spinnaker](#)
- ▶ [Monitoring Distributed Systems: The Four Golden Signals](#)





QUESTIONS?

