

13th Symposium on Software Performance 2022

Cloud-Native Scalability Benchmarking with Theodolite

Sören Henning, Benedikt Wetzel, Wilhelm Hasselbring



Scalability Benchmarking of Cloud-Native Applications

Scalability is a main driver for adopting cloud-native applications and microservices

(Knoche & Hasselbring 2019; Soldani et al. 2018; Kratzke & Quint 2017)

Scalability Benchmarking of Cloud-Native Applications

 Scalability is a main driver for adopting cloud-native applications and microservices

(Knoche & Hasselbring 2019; Soldani et al. 2018; Kratzke & Quint 2017)

 Engineers and researchers use benchmarking to evaluate and compare quality of frameworks, configuration, etc.

(Kounev et al. 2020; Hasselbring 2021)

Scalability Benchmarking of Cloud-Native Applications

Scalability is a main driver for adopting cloud-native applications and microservices

(Knoche & Hasselbring 2019; Soldani et al. 2018; Kratzke & Quint 2017)

 Engineers and researchers use benchmarking to evaluate and compare quality of frameworks, configuration, etc.
 (Kounev et al. 2020; Hasselbring 2021)

 However, no commonly used method for benchmarking scalability of cloud-native applications

(Henning & Hasselbring 2022)

Goal:

Make Scalability Benchmarking more Usable and Reproducible!

The Theodolite Scalability Benchmarking Framework





Creating benchmarks based on existing **Kubernetes** resource files



Defining SLO based on existing **Prometheus** metrics via **PromQL** queries



Kubernetes Operator for declarative definition of benchmarks and executions

In Search of Microservice Show Case...

TeaStore













Green Tea

Black Tea

Categories

Pure black tea and blends

Green Tea

From China and Japan

Herbal Tea

Helps when you feel sick

Rooibos

In many variations

White Tea

If green tea doesn't agree with you

Black Tea

Pure black tea and blends

Green Tea

From China and Japan

Herbal Tea

Helps when you feel sick

Rooibos

In many variations

White Tea

If green tea doesn't agree with you

Black Tea

Pure black tea and blends

Green Tea

From China and Japan

Herbal Tea

Helps when you feel sick

Rooibos

In many variations

White Tea

If green tea doesn't agree with you



Price: 66,70 \$

Great Green Tea: Sencha (loose)

Add to Cart



Sencha (15 bags)

Price: 3,48 \$

Great Green Tea: Sencha (15 bags)

Add to Cart



Sencha (25 bags)

Price: 109,35 \$

Great Green Tea: Sencha (25 bags)

Add to Cart



Earl Grey Green (loose)

Logout

Price: 80,63 \$

Great Green Tea: Earl Grey Gree...

Add to Cart



Earl Grey Green (15 bags)

Price: 48,39 \$

Great Green Tea: Earl Grey Gree...

Add to Cart



Earl Grey Green (25 bags)

Price: 82,44 \$

Great Green Tea: Earl Grey Gree ...

Add to Cart



Matcha 30 g

Price: 10,18 \$

Great Green Tea: Matcha 30 g

Add to Cart



Matcha 50 g

Price: 5,65 \$

Great Green Tea: Matcha 50 g

Add to Cart



Matcha 100 g

Price: 118,59 \$

Great Green Tea: Matcha 100 g

Add to Cart



Gunpowder Tea (loose)

Price: 32,74 \$

Great Green Tea: Gunpowder Tea..

Add to Cart



Add to Cart





Gunpowder Tea (25 bags)

Price: 71,50 \$

Great Green Tea: Gunpowder Tea...

Add to Cart



Sencha (loose), v1

Price: 76,58 \$

Great Green Tea: Sencha (loose)

Add to Cart



Sencha (15 bags), v1

Price: 81,23 \$

Great Green Tea: Sencha (15 bags)

Add to Cart



Sencha (25 bags), v1

Price: 11.93 \$

Great Green Tea: Sencha (25 bags)

Add to Cart



Earl Grey Green (loose), v1

Price: 86,99 \$

Great Green Tea: Earl Grey Gree ...

Add to Cart



Earl Grey Green (15 bags), v1

Price: 74,83 \$

Great Green Tea: Earl Grey Gree...

Add to Cart



Earl Grey Green (25 bags), v1

Price: 69,06 \$

Add to Cart



Matcha 30 g, v1 Price: 85,38 \$

Great Green Tea: Matcha 30 g

Add to Cart



Matcha 50 g, v1

Great Green Tea: Matcha 50 g

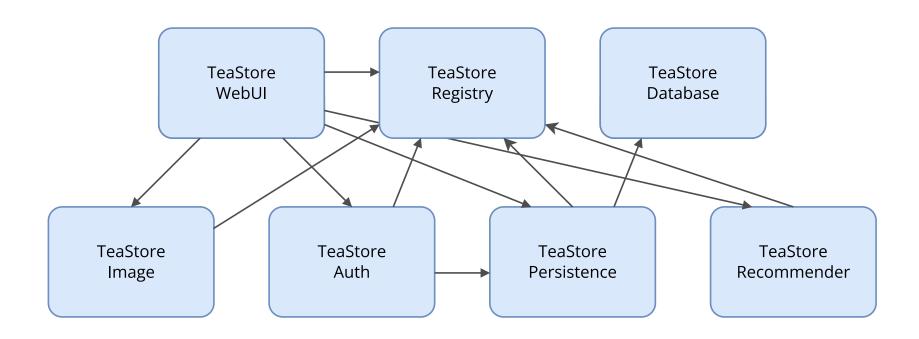
Add to Cart

2 3 4 5 6 next

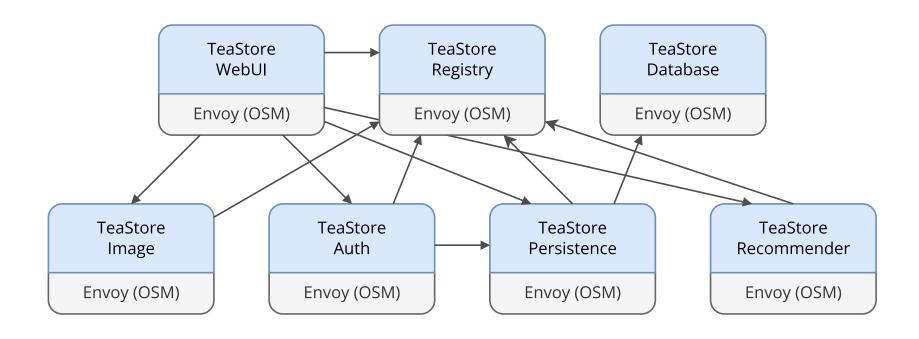
20 v products per page



TeaStore Architecture



TeaStore Architecture

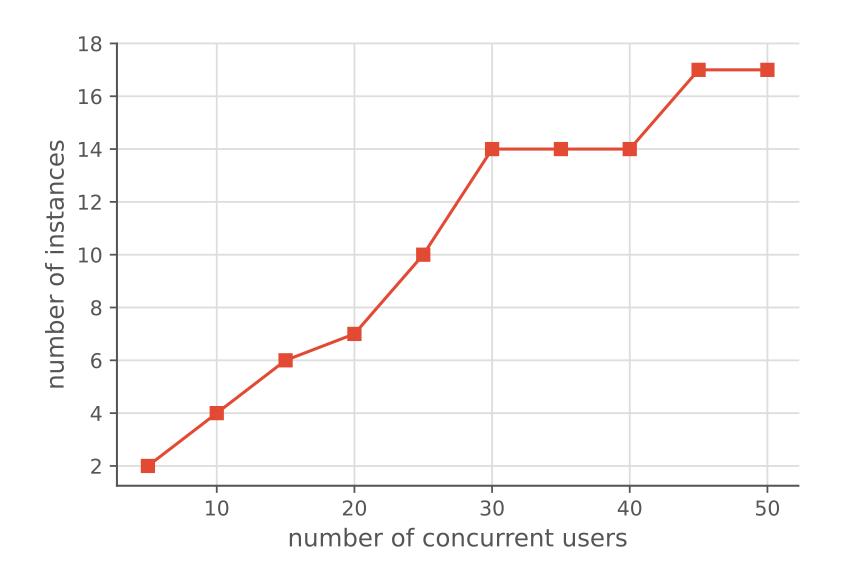






Open Service Mesh

Desired Outcome



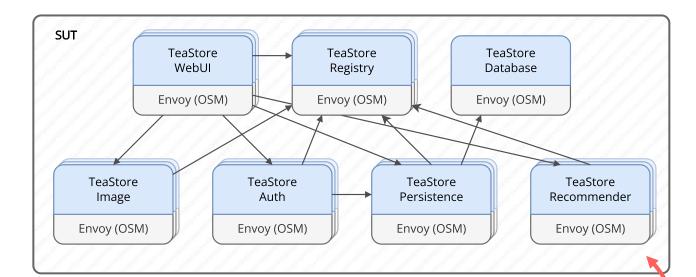
Designing a Scalability Benchmark for the TeaStore

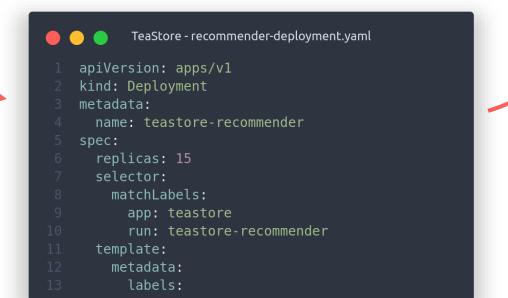
```
TeaStore - benchmark.yaml
    apiVersion: theodolite.rocks/v1beta1
    kind: benchmark
    metadata:
      name: teastore
    spec:
      waitForResourcesEnabled: true
       sut:
        resources:
        configMap:
            name: teastore-deployment
            files:
            - teastore-auth-deployment.yaml
            - teastore-auth-service.yaml
            - teastore-db-deployment.yaml
            - teastore-db-service.yaml
            - teastore-image-deployment.yaml
            - teastore-image-service.yaml
            - teastore-persistence-deployment.yaml
            - teastore-persistence-service.yaml
            - teastore-recommender-deployment.yaml
            - teastore-recommender-service.yaml
            - teastore-registry-deployment.yaml
            - teastore-registry-service.yaml
            - teastore-webui-deployment.yaml
            - teastore-webui-nodeport.yaml
       loadGenerator:
        resources:
         configMap:
            name: teactore imeter denloyment
```

```
TeaStore - benchmark.yaml
```

```
apiVersion: theodolite.rocks/v1beta1
    kind: benchmark
    metadata:
       name: teastore
    spec:
       waitForResourcesEnabled true
       sut:
         resources:
         configMap:
            name: teastore-deployment
            files:
            - teastore-auth-deployment.yaml
            - teastore-auth-service.yaml
            - teastore-db-deployment.yaml
            - teastore-db-service.yaml
            - teastore-image-deployment.yaml
            - teastore-image-service.yaml
            - teastore-persistence-deployment.yaml
            - teastore-persistence-service.yaml
            - teastore-recommender-deployment.yaml
            - teastore-recommender-service.yaml
            - teastore-registry-deployment.yaml
            - teastore-registry-service.yaml
            - teastore-webui-deployment.yaml
            - teastore-webui-nodeport.yaml
        resources:
         - configMap:
            name: teastore-jmeter-deployment
            files:
            - jmeter.yaml
       resourceTypes:
      - typeName: Instances
        patchers:
        - type: ReplicaPatcher
          resource: teastore-auth-deployment.yaml
        - type: ReplicaPatcher
           resource: teastore-image-deployment.yaml
         - type: ReplicaPatcher
```

System under Test (SUT)





13

TeaStore - benchmark.yaml apiVersion: theodolite.rocks/v1beta1 kind: benchmark metadata: name: teastore spec: waitForResourcesEnabled: true sut: resources: configMap: name: teastore-deployment files: - teastore-auth-deployment.yaml - teastore-auth-service.yaml - teastore-db-deployment.yaml - teastore-db-service.yaml - teastore-image-deployment.yaml - teastore-image-service.yaml - teastore-persistence-deployment.yaml - teastore-persistence-service.yaml - teastore-recommender-deployment.yaml - teastore-recommender-service.yaml - teastore-registry-deployment.yaml - teastore-registry-service.yaml - teastore-webui-deployment.yaml loadGenerator: resources: configMap: name: teastore-jmeter-deployment files: - jmeter.yaml resourceTypes. - typeName: Instances patchers: - type: ReplicaPatcher resource: teastore-auth-deployment.yaml - type: ReplicaPatcher

resource: teastore-image-deployment.yaml

- type: ReplicaPatcher

Load Generator

Load Generator JMeter Envoy (OSM)

TeaStore - jmeter.yaml

```
apiVersion: apps/v1
kind: Deployment
metadata:
  name: jmeter
  annotations:
    openservicemesh.io/sidecar-injection: enabled
spec:
  replicas: 1
      containers:
      - name: jmeter
        image: justb4/jmeter:5.4
        command: [ "/bin/bash", "-c", "--" ]
        args:
          java -jar bin/ApacheJMeter.jar \
          -t profile/teastore browse nogui.jmx \
          -Jhostname teastore-webui -Jport 8080 \
          -JnumUser $NUM USERS -JrampUp 1 \
          -l mylogfile.log -n
        env:
        - name: NUM USERS
          value: "20"
        volumeMounts:
```

```
name: teastore-deployment
      files:
     - teastore-auth-deployment.yaml
     - teastore-auth-service.yaml
      - teastore-db-deployment.yaml
     - teastore-db-service.yaml
     - teastore-image-deployment.yaml
     - teastore-image-service.yaml
     - teastore-persistence-deployment.yaml
     - teastore-persistence-service.yaml
     - teastore-recommender-deployment.yaml
     - teastore-recommender-service.yaml
     - teastore-registry-deployment.yaml
     - teastore-registry-service.yaml
     - teastore-webui-deployment.yaml
     - teastore-webui-nodeport.yaml
loadGenerator:
 resources:
 configMap:
     name: teastore-jmeter-deployment
     files:
      - imeter vaml
resourceTypes:
- typeName: Instances
 patchers:
 - type: ReplicaPatcher
   resource: teastore-auth-deployment.yaml
 - type: ReplicaPatcher
   resource: teastore-image-deployment.yaml
 - type: ReplicaPatcher
   resource: teastore-persistence-deployment.yan<mark>l</mark>
 - type: ReplicaPatcher
   resource: teastore-recommender-deployment.yan
 - type: ReplicaPatcher
   resource: teastore-webui-deployment.yaml
typeName: "PodResources"
 patchers:
 - type: ResourceLimitPatcher
   resource: teastore-auth-deployment.yaml
   properties:
     container: teastore-auth
     limitedResource: cpu
     format: m
 - type: ResourceLimitPatcher
```

Resource Dimension

(1) Horizontal Scalability

```
TeaStore - recommender-deployment.yaml
apiVersion: apps/v1
kind: Deployment
metadata:
  name: teastore-recommender
spec:
  replicas: 15
  selector:
    matchLabels:
      app: teastore
      run: teastore-recommender
  template:
    metadata:
      labels:
        app: teastore
        run: teastore-recommender
    spec:
      terminationGracePeriodSeconds: 0
      containers:
      - name: teastore-recommender
        image: descartesresearch/teastore-reco
mmender:development
        imagePullPolicy: Always
        ports:
        - containerPort: 8080
```

```
resource: teastore-recommender-deployment.yaml
 - type: ReplicaPatcher
   resource: teastore-webui-deployment vaml
- typeName: "PodResources"
 patchers:
 - type: ResourceLimitPatcher
   resource: teastore-auth-deployment.yaml
   properties:
   properties:
     container: teastore-auth
     limitedResource: memory
     format: Mi
     factor: 2
 - type: ResourceLimitPatcher
   resource: teastore-image-deployment.yaml
   properties:
     container: teastore-image
     limitedResource: cpu
     format: m
 - type: ResourceLimitPatcher
   resource: teastore-image-deployment.yaml
   properties:
     container: teastore-image
     limitedResource: memory
     format: Mi
     factor: 2
 - type: ResourceLimitPatcher
   resource: teastore-recommender-deployment.yaml
   properties:
     container: teastore-recommender
     limitedResource: cpu
     format: m
 - type: ResourceLimitPatcher
   resource: teastore-recommender-deployment.yaml
   properties:
     container: teastore-recommender
     limitedResource: memory
     format: Mi
     factor: 2
 - type: ResourceLimitPatcher
   resource: teastore-webui-deployment.yaml
   properties:
```

- type: ReplicaPatcher

Resource Dimension

(2) Vertical Scalability

```
TeaStore - recommender-deployment.yaml
apiVersion: apps/v1
kind: Deployment
metadata:
  name: teastore-recommender
spec:
           TaltureInresnotu. 5
          httpGet:
            path: "/tools.descartes.teastore.r
ecommender/rest/train/isready"
            port: 8080
            scheme: HTTP
         resources:
          limits:
            memory: 1000Mi
            cpu: 500m
      nodeSelector:
        env: dev
```

```
properties:
        container: teastore-recommender
        limitedResource: memory
        format: Mi
        factor: 2
    - type: ResourceLimitPatcher
      resource: teastore-webui-deployment.yaml
      properties:
        container: teastore-webui
        limitedResource: cpu
        format: m
    - type: ResourceLimitPatcher
      resource: teastore-webui-deployment.yaml
      properties:
        container: teastore-webui
        limitedResource: memory
        format: Mi
  loadTypes:
  - typeName: NumUsers
    patchers:
    - type: EnvVarPatcher
      resource: jmeter.yaml
      properties:
        variableName: NUM USERS
    - sloType: generic
     name: uiLatency
      prometheusUrl: "http://prometheus-operated:9090"
     offset: 0
      properties:
        externalSloUrl: "http://localhost:8082"
        promQLQuery: "histogram quantile(0.95,sum(irat
e(osm request duration ms bucket{destination name='tea
store webui'}[1m])) by (le, destination name))"
        warmup: 600 # in seconds
        queryAggregation: max
        repetitionAggregation: median
        operator: lte
        threshold: 200
```

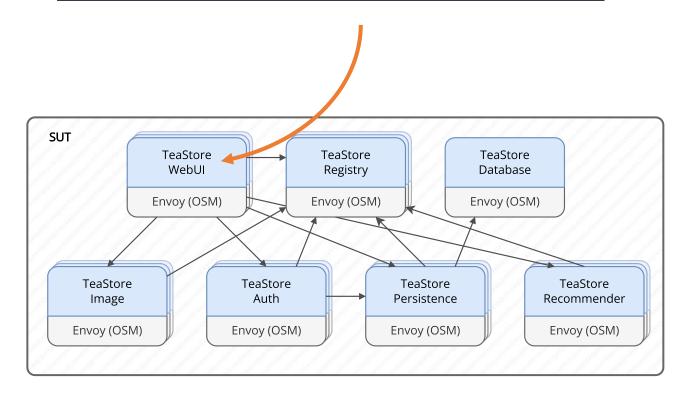
Load Dimension

```
TeaStore - jmeter.yaml
    apiVersion: apps/v1
   kind: Deployment
   metadata:
      name: jmeter
      annotations:
        openservicemesh.io/sidecar-injection: enabled
   spec:
      replicas: 1
          containers:
          - name: jmeter
            image: justb4/jmeter:5.4
            command: [ "/bin/bash", "-c", "--" ]
            args:
              java -jar bin/ApacheJMeter.jar \
              -t profile/teastore browse nogui.jmx \
              -Jhostname teastore-webui -Jport 8080 \
              -JnumUser $NUM USERS -JrampUp 1 \
              -l mylogfile.log -n
            env:
            - name: NUM USERS
              value: "20"
            volumeMounts:
            - name: jmeter-profile
              mountPath: /opt/apache-jmeter-5.4.3/profile
            ports:
            - containerPort: 50000
            - containerPort: 60000
            - containerPort: 1099
          volumes•
```

```
properties:
       container: teastore-recommender
       limitedResource: memory
       format: Mi
       factor: 2
    - type: ResourceLimitPatcher
      resource: teastore-webui-deployment.yaml
     properties:
       container: teastore-webui
       limitedResource: cpu
        format: m
    - type: ResourceLimitPatcher
      resource: teastore-webui-deployment.yaml
     properties:
       container: teastore-webui
       limitedResource: memory
       format: Mi
       factor: 2
 loadTypes:
 - typeName: NumUsers
   patchers:
    - type: EnvVarPatcher
     resource: jmeter.yaml
     properties:
       container: jmeter
        variableName: NUM USERS
    sloType: generic
     name: uiLatency
     prometheusUrl: "http://prometheus-operated:9090"
     offset: 0
     properties:
       externalSloUrl: "http://localhost:8082"
       promQLQuery: "histogram quantile(0.95,sum(irat
e(osm request duration ms bucket{destination name='tea
store webui'}[1m])) by (le, destination name))"
       warmup: 600 # in seconds
       queryAggregation: max
       repetitionAggregation: median
       operator: lte
       threshold: 200
```

SLO

p95 latency <= 200 ms?



TeaStore - benchmark.yaml

```
apiVersion: theodolite.rocks/v1beta1
kind: benchmark
metadata:
  name: teastore
spec:
  waitForResourcesEnabled: true
  sut:
    resources:
    configMap:
        name: teastore-deployment
        files:
        - teastore-auth-deployment.yaml
        - teastore-auth-service.yaml
        - teastore-db-deployment.yaml
        - teastore-db-service.yaml
        - teastore-image-deployment.yaml
        - teastore-image-service.yaml
        - teastore-persistence-deployment.yaml
        - teastore-persistence-service.yaml
        - teastore-recommender-deployment.yaml
        - teastore-recommender-service.yaml
        - teastore-registry-deployment.yaml
        - teastore-registry-service.yaml
        - teastore-webui-deployment.yaml
        - teastore-webui-nodeport.yaml
  loadGenerator:
    resources:
    configMap:
        name: teastore-jmeter-deployment
        files:
        - jmeter.yaml
  resourceTypes:
  - typeName: Instances
    patchers:
    - type: ReplicaPatcher
      resource: teastore-auth-deployment.yaml
    - type: ReplicaPatcher
      resource: teastore-image-deployment.yaml
    - type: ReplicaPatcher
```

Deployment

```
TeaStore - bash

$ kubectl apply -f benchmark.yaml
benchmark.theodolite.rocks/teastore created
```

```
TeaStore - benchmark.yaml
```

```
apiVersion: theodolite.rocks/v1beta1
kind: benchmark
metadata:
  name: teastore
spec:
  waitForResourcesEnabled: true
  sut:
    resources:
    configMap:
        name: teastore-deployment
        files:
        - teastore-auth-deployment.yaml
        - teastore-auth-service.yaml
        - teastore-db-deployment.yaml
        - teastore-db-service.yaml
        - teastore-image-deployment.yaml
        - teastore-image-service.yaml
        - teastore-persistence-deployment.yaml
        - teastore-persistence-service.yaml
        - teastore-recommender-deployment.yaml
        - teastore-recommender-service.yaml
        - teastore-registry-deployment.yaml
        - teastore-registry-service.yaml
        - teastore-webui-deployment.yaml
        - teastore-webui-nodeport.yaml
  loadGenerator:
    resources:
    configMap:
        name: teastore-jmeter-deployment
        files:
        - jmeter.yaml
  resourceTypes:
  - typeName: Instances
    patchers:
    - type: ReplicaPatcher
      resource: teastore-auth-deployment.yaml
    - type: ReplicaPatcher
      resource: teastore-image-deployment.yaml
    - type: ReplicaPatcher
```

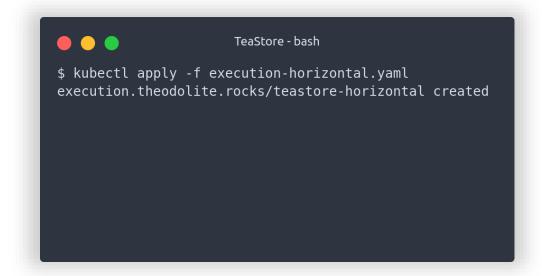
Deployment

```
TeaStore-bash

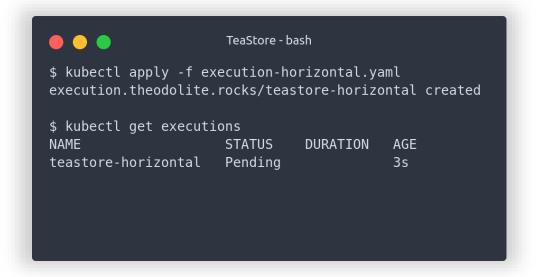
$ kubectl apply -f benchmark.yaml
benchmark.theodolite.rocks/teastore created

$ kubectl get benchmarks
NAME AGE STATUS
teastore 25s Ready
```

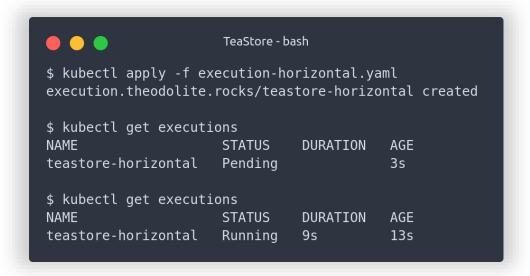
```
TeaStore - execution-horizontal.yaml
   apiVersion: theodolite.rocks/v1beta1
   kind: execution
   metadata:
     name: teastore-horizontal
5 spec:
     benchmark: teastore
     load:
       loadType: NumUsers
       loadValues: [5, 10, 15, 20, 25, 30, 35, 40, 45, 50]
     resources:
       resourceType: Instances
       resourceValues: [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11,
   12, 13, 14, 15, 16, 17, 18, 19, 20]
       - name: uiLatency
         properties:
          warmup: 600 # in seconds
           threshold: 200
     execution:
       strategy:
       name: RestrictionSearch
       restrictions:
        - LowerBound
         searchStrategy: LinearSearch
       duration: 1200 # in seconds
       repetitions: 1
     configOverrides: []
```



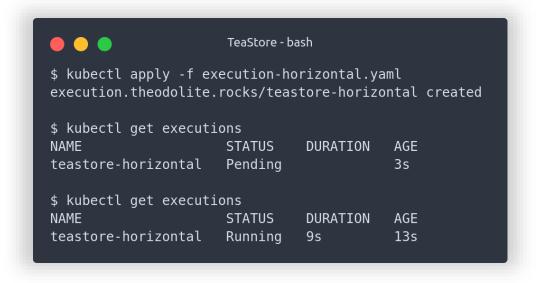
```
TeaStore - execution-horizontal.yaml
   apiVersion: theodolite.rocks/v1beta1
   kind: execution
   metadata:
     name: teastore-horizontal
5 spec:
     benchmark: teastore
     load:
      loadType: NumUsers
      loadValues: [5, 10, 15, 20, 25, 30, 35, 40, 45, 50]
     resources:
       resourceType: Instances
      resourceValues: [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11,
   12, 13, 14, 15, 16, 17, 18, 19, 20]
       - name: uiLatency
         properties:
          warmup: 600 # in seconds
           threshold: 200
     execution:
       strategy:
       name: RestrictionSearch
       restrictions:
        - LowerBound
         searchStrategy: LinearSearch
       duration: 1200 # in seconds
       repetitions: 1
     configOverrides: []
```



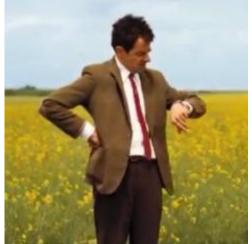
```
TeaStore - execution-horizontal.yaml
   apiVersion: theodolite.rocks/v1beta1
   kind: execution
   metadata:
     name: teastore-horizontal
5 spec:
     benchmark: teastore
     load:
       loadType: NumUsers
      loadValues: [5, 10, 15, 20, 25, 30, 35, 40, 45, 50]
     resources:
       resourceType: Instances
      resourceValues: [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11,
   12, 13, 14, 15, 16, 17, 18, 19, 20]
       name: uiLatency
         properties:
          warmup: 600 # in seconds
           threshold: 200
     execution:
       strategy:
       name: RestrictionSearch
       restrictions:
        - LowerBound
         searchStrategy: LinearSearch
       duration: 1200 # in seconds
       repetitions: 1
     configOverrides: []
```

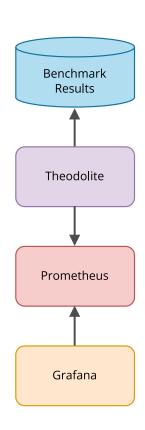


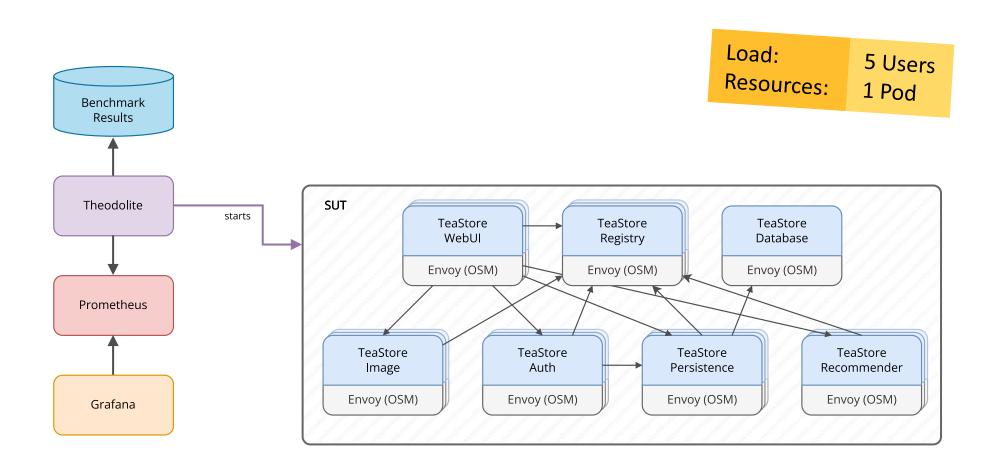
```
TeaStore - execution-horizontal.yaml
apiVersion: theodolite.rocks/v1beta1
kind: execution
metadata:
  name: teastore-horizontal
spec:
  benchmark: teastore
  load:
    loadType: NumUsers
    loadValues: [5, 10, 15, 20, 25, 30, 35, 40, 45, 50]
  resources:
    resourceType: Instances
    resourceValues: [1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11,
12, 13, 14, 15, 16, 17, 18, 19, 20]
    - name: uiLatency
      properties:
       warmup: 600 # in seconds
        threshold: 200
  execution:
    strategy:
    name: RestrictionSearch
    restrictions:
     - LowerBound
      searchStrategy: LinearSearch
    duration: 1200 # in seconds
    repetitions: 1
  configOverrides: []
```

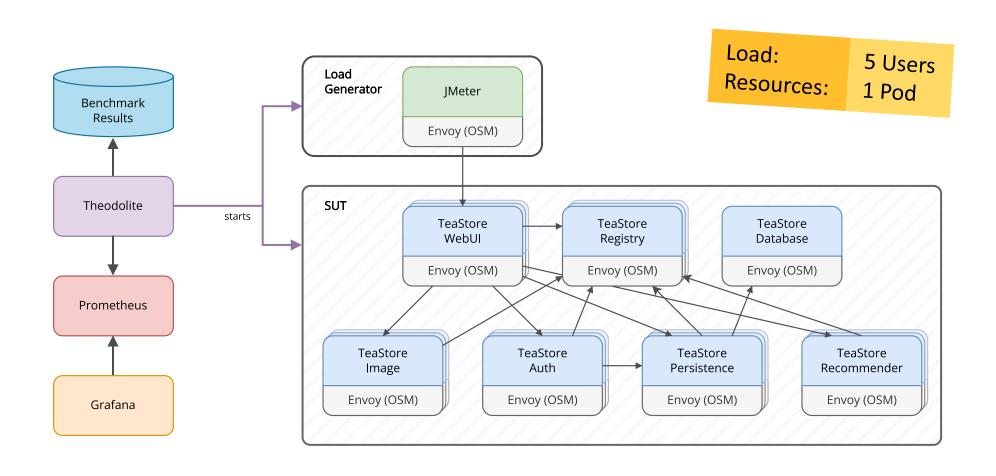


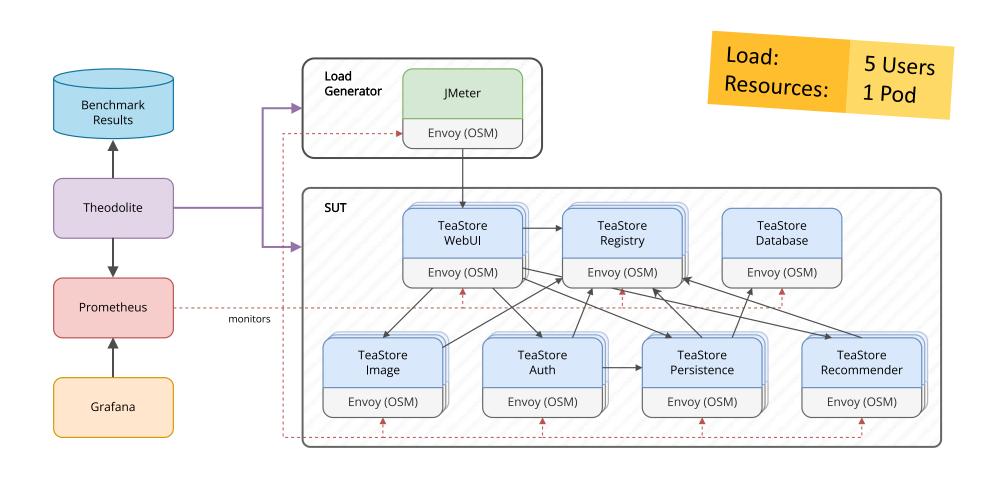


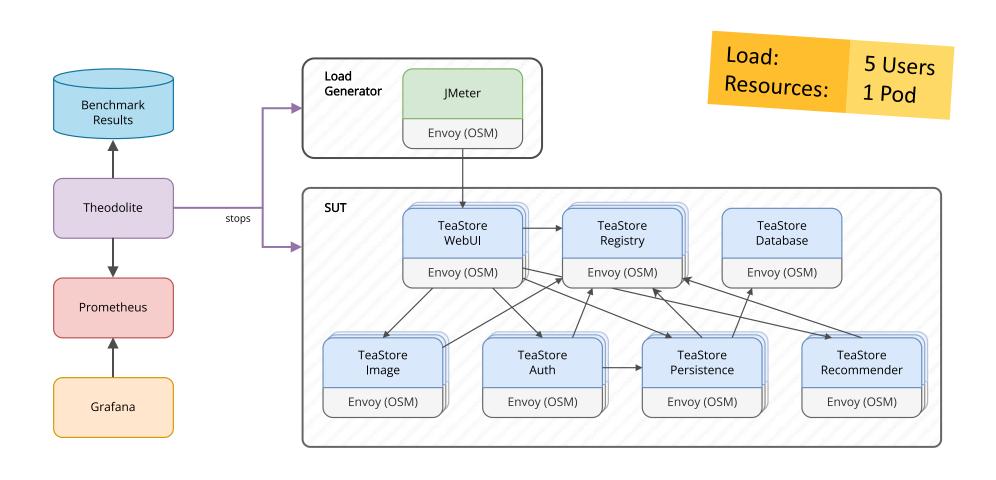


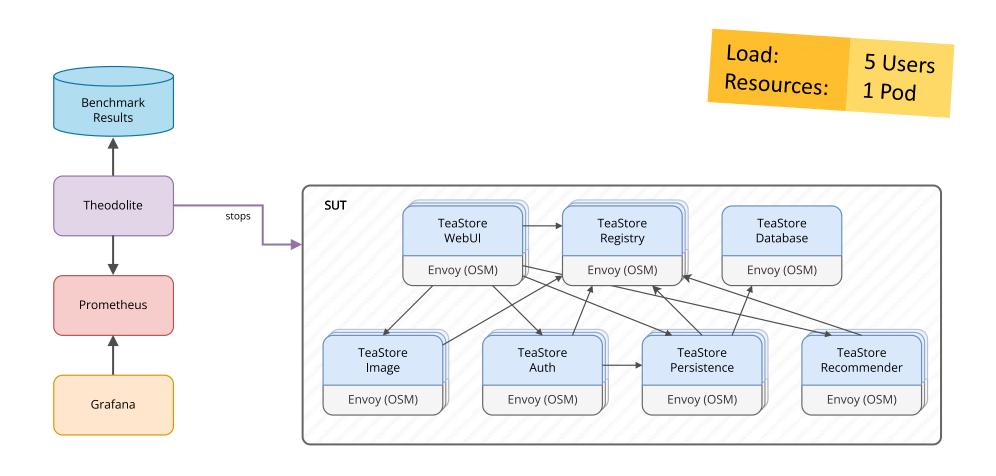


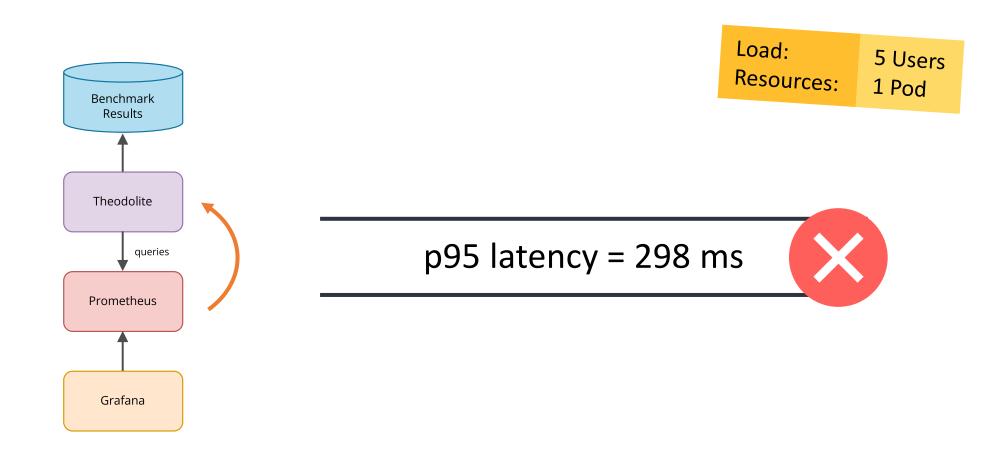


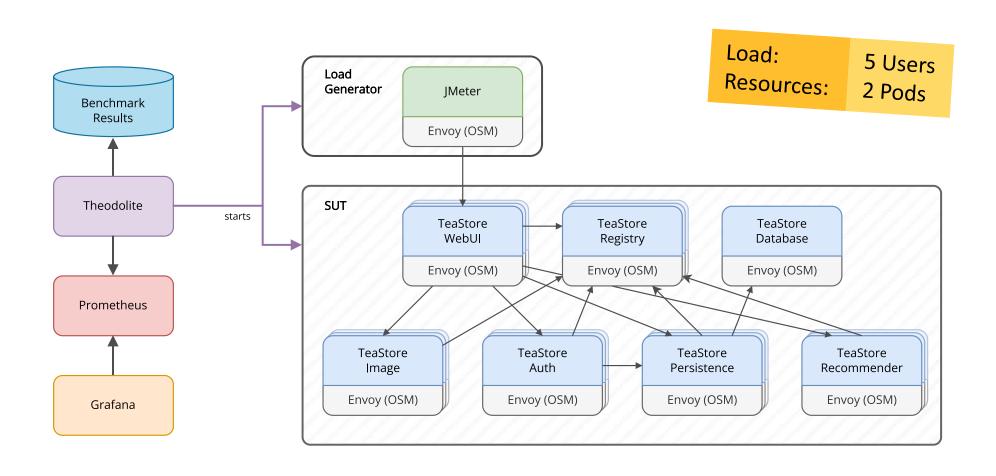


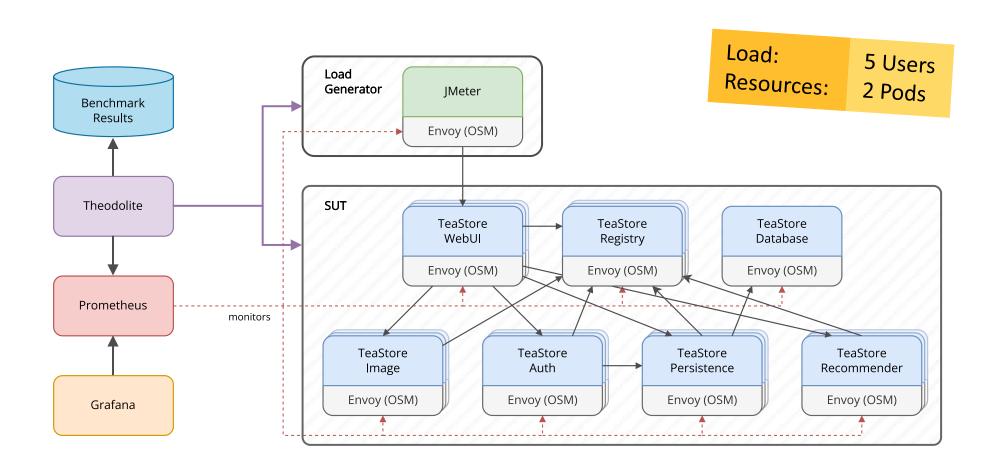


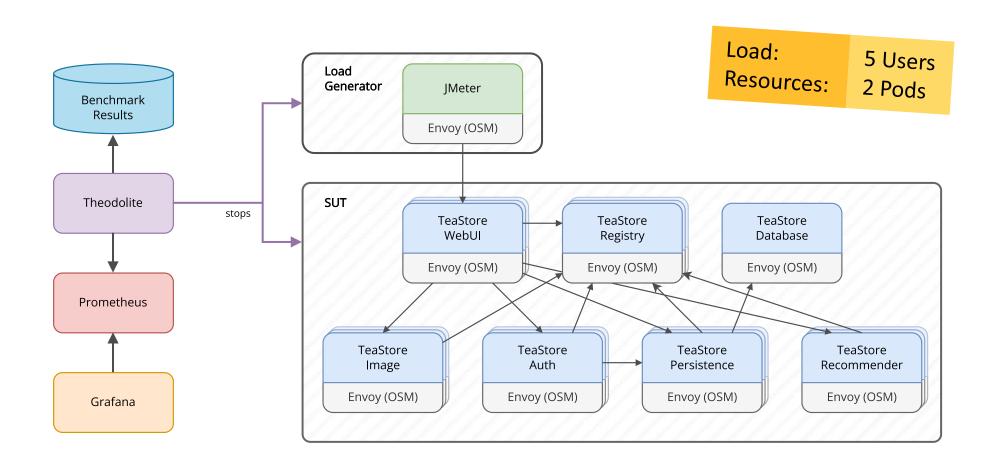


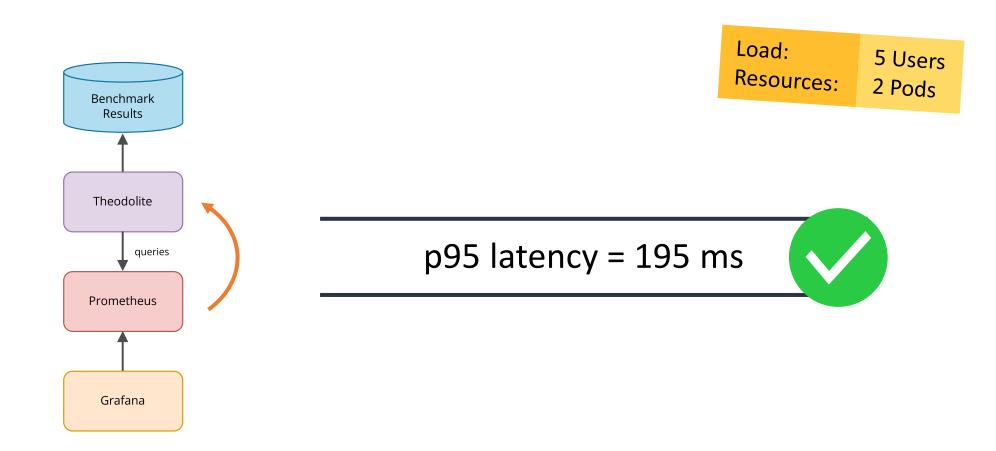


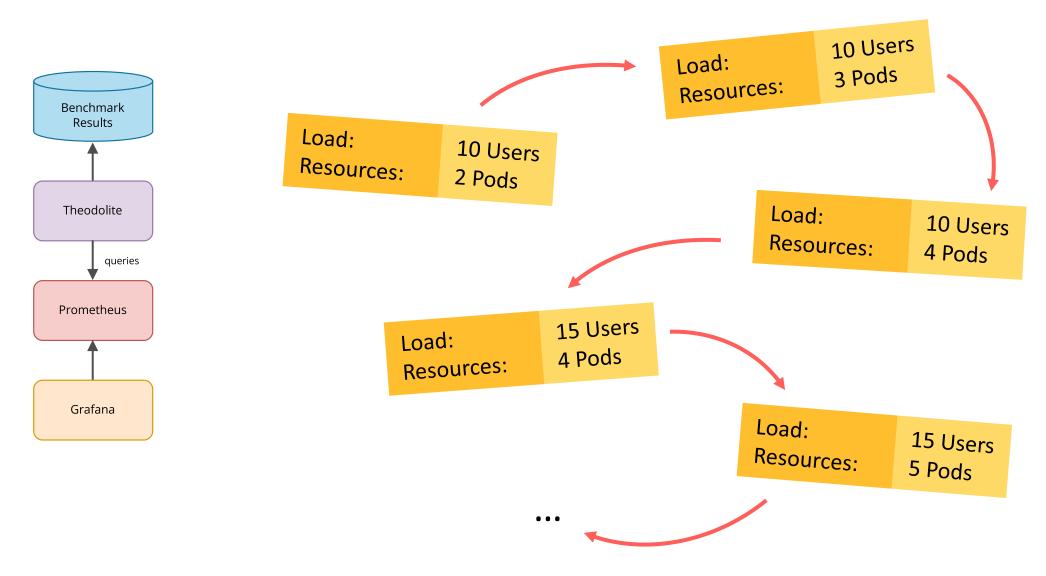




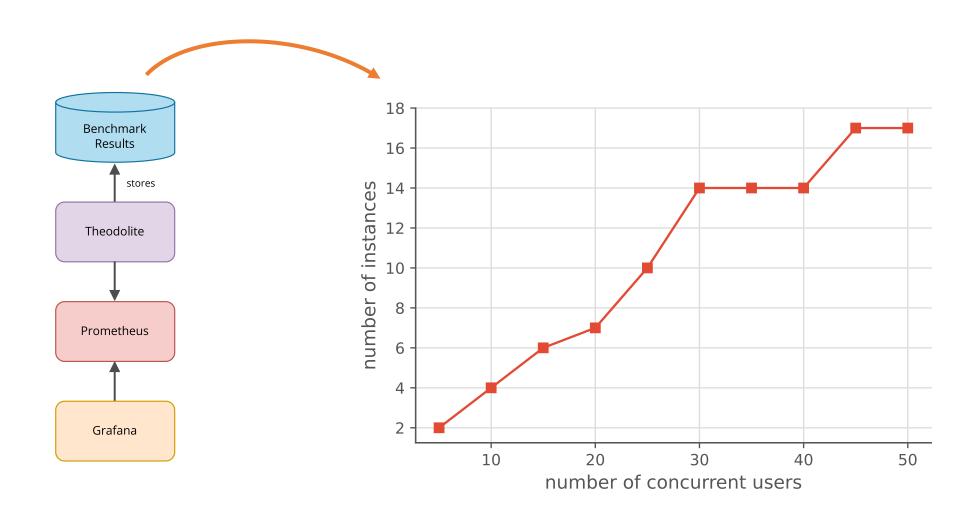






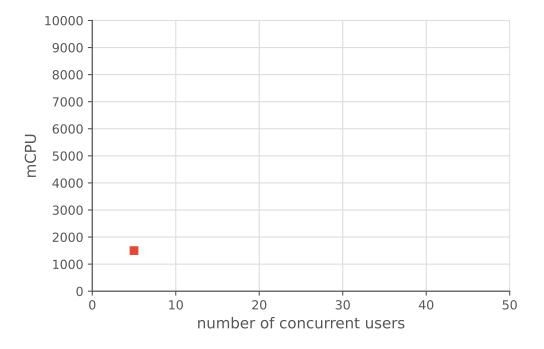


...and the result is:



Benchmarking Vertical Scalability

```
TeaStore - execution-vertical.yaml
apiVersion: theodolite.rocks/v1beta1
kind: execution
metadata:
  name: teastore-vertical
spec:
  benchmark: teastore
  load:
   loadType: NumUsers
    loadValues: [5, 10, 15, 20, 25, 30, 35, 40, 45, 50]
  resources:
    resourceType: PodResources
    resourceValues: [500, 1000, 1500, 2000, 2500, 3000,
3500, 4000, 4500, 5000, 5500, 6000, 6500, 7000, 7500, 8
000, 8500, 9000, 9500, 10000]
    - name: uiLatency
      properties:
       warmup: 600 # in seconds
       threshold: 200
  execution:
   strategy:
      name: RestrictionSearch
      restrictions:
      - LowerBound
     searchStrategy: LinearSearch
    duration: 1200 # in seconds
    repetitions: 1
  configOverrides: []
```

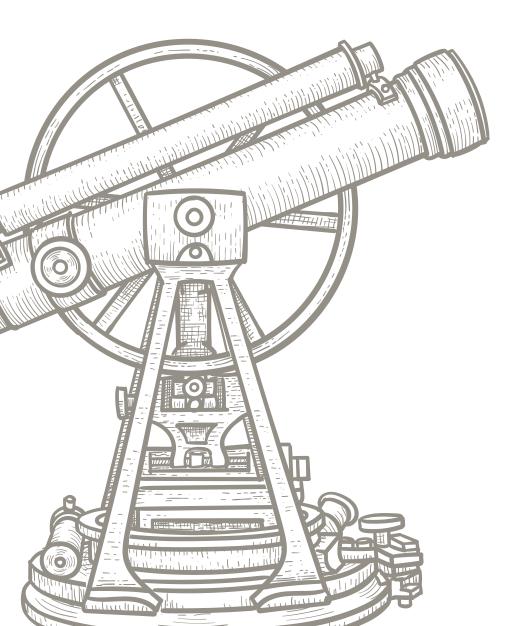


Conclusions & Lessons Learned



- Favor open workload over closed workload models
- Finding **good SLOs** is hard
- Cloud-native-ness of the TeaStore could be improved
- Interested? Theodolite's TeaStore benchmark is now ready to use!

Getting Started



Theodolite



