TOPIC

Operate Kubernetes Workloads:

Extend the platform with the operator pattern!

Paolo Patierno, Principal Software Engineer @RedHat

KIRATECH®

WE DEVOPS IT

#CodeGen2021

@cloudgen_verona

Who am I?

```
apiVersion: v1
kind: PrincipalSoftwareEngineer
metadata:
    name: Paolo Patierno
    namespace: Red Hat, Messaging & Data Streaming
    labels:
        cncf/maintainer: Strimzi
        eclipse/committer: Vert.x, Hono & Paho
        microsoft/mvp: Azure
    annotations:
        family: dad of two, husband of one
        sports: running, swimming, motogp, vr46, formula1, ferrari, ssc napoli
        community: cncf napoli, devday
spec:
    replicas: 1
    containers:
    - image: patiernohub.io/paolo:latest
```

@ppatierno

Kubernetes

“ A system for …”

“ … automating deployment …”

“ … scaling …”

“ … management …”

“ … of containerized applications …”

“ It’s like a Linux kernel … but for distributed systems”

Kubernetes

Container scheduling

Self healing

Secret & configuration management

Service discovery

Horizontal scaling

Load balancing

Storage orchestration

Automated rollout/rollback

Batch execution

Kubernetes workloads

“

How does Kubernetes handle scaling, rollout, batch execution and so on?

Kubernetes workloads

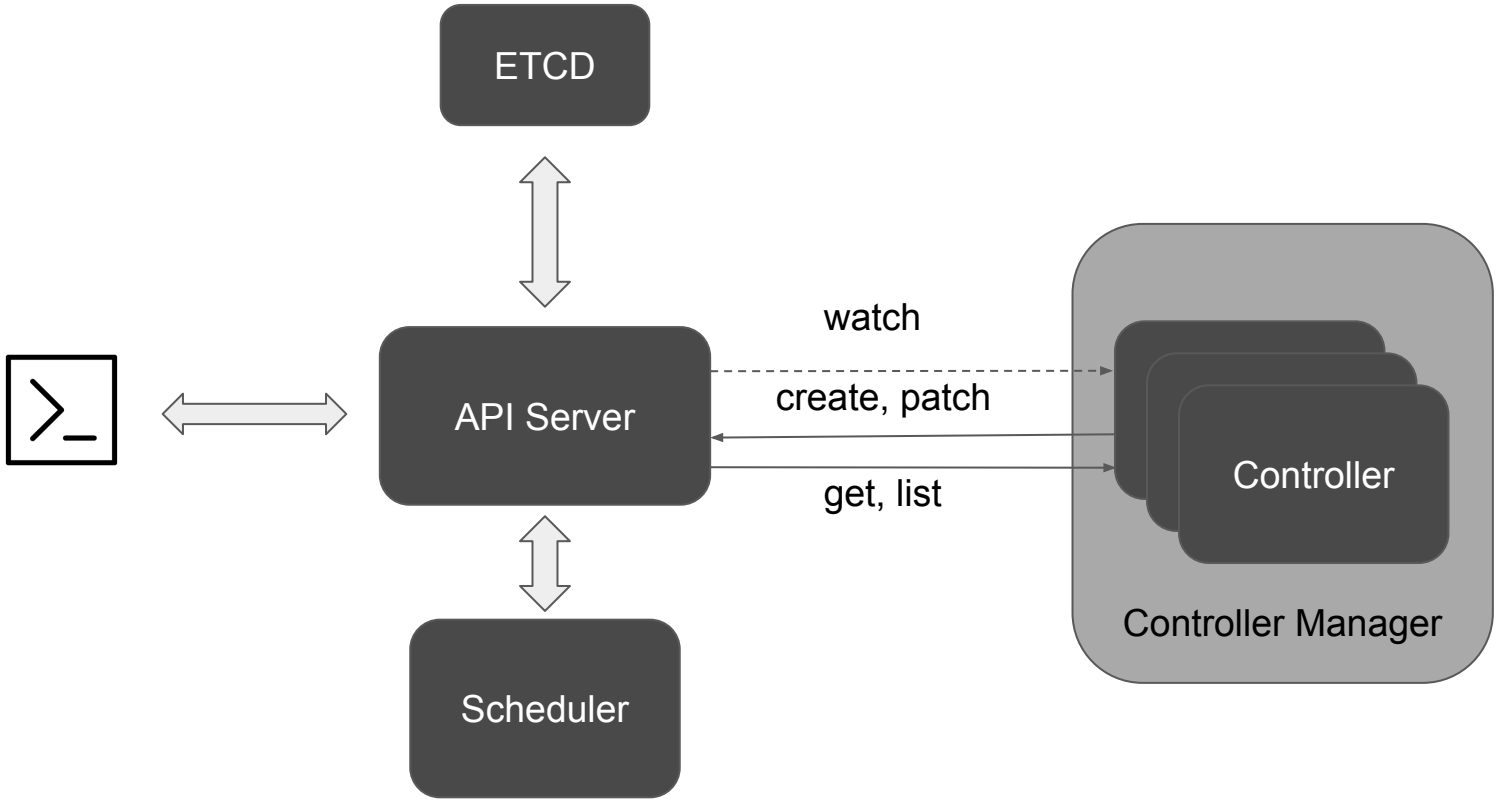
- Don’t use Pod(s) … let’s use something more sophisticated!
- ReplicaSet
  - Guarantees a specific number of running replicas
  - Spins pods, based on a template, if there are not enough
  - Deletes pods if too many match the selector
- Deployment
  - It’s based on ReplicaSet (used to run replicas)
  - Adds extra layer for rollout and rollback
- … and more with StatefulSet, Job, DaemonSet ...

How does Kubernetes do the trick? What’s behind these special resources?

Kubernetes controllers

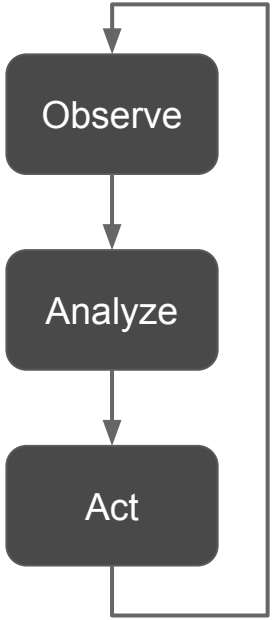
- A controller …
  - … tracks a Kubernetes resource/object
  - … makes sure that the object specification, desired state, matches the current state on cluster
  - … interacts with the API server
  - … acts in a control loop
- Built-in controllers
  - ReplicaSet
  - Deployment
  - … the Kubernetes controller manager runs them

Kubernetes controllers



Controller loop

- Observe
  - Watch for resource/object creation or changes
- Analyze
  - Check that the resource/object desired state (“spec”) reflects the current state on the cluster
- Act
  - Makes the needed changes



ReplicaSet controller


```yaml
apiVersion: apps/v1
kind: ReplicaSet
metadata:
  name: my-replicaset
spec:
  replicas: 3
  selector:
    matchLabels:
      app: my-app
  template:
    metadata:
      labels:
        app: my-app
    spec:
      containers:
      - name: my-application
        image: quay.io/ppatierno/my-application:latest
```

my-replicaset-bf5zv

my-replicaset-1tf5a

my-replicaset-gb65f

ReplicaSet controller


```yaml
apiVersion: v1
kind: Pod
metadata:
  name: my-pod-1
  labels:
    app: my-app
spec:
  containers:
  - name: my-app
    image: quay.
```

```yaml
apiVersion: v1
kind: Pod
metadata:
  name: my-pod-2
  labels:
    app: my-app
spec:
  containers:
  - name: my-application
    image: quay.io/ppatierno/my-application:latest
```

my-replicaset-bf5zv

my-replicaset-1tf5a

my-replicaset-gb65f

~~my-pod-1~~

~~my-pod-2~~

ReplicaSet controller

my-pod-1

my-pod-2

```yaml
apiVersion: v1
kind: Pod
metadata:
  name: my-pod-1
  labels:
    app: my-app
spec:
  containers:
  - name: my-app
    image: quay.
```

```yaml
apiVersion: v1
kind: Pod
metadata:
  name: my-pod-2
  labels:
    app: my-app
spec:
  containers:
  - name: my-application
    image: quay.io/ppatierno/my-application:latest
```

ReplicaSet controller


```yaml
apiVersion: apps/v1
kind: ReplicaSet
metadata:
  name: my-replicaset
spec:
  replicas: 3
  selector:
    matchLabels:
      app: my-app
  template:
    metadata:
      labels:
        app: my-app
    spec:
      containers:
      - name: my-application
        image: quay.io/ppatierno/my-application:latest
```

my-pod-1

my-pod-2

my-replicaset-65rt3

“

Humans build complex applications with these “bricks”

Human operating Kubernetes workloads

- Human operator knows ...
  - … about a complex application or service
  - … how it behaves internally
  - … how to deploy it
  - … how to react and fix issues with it

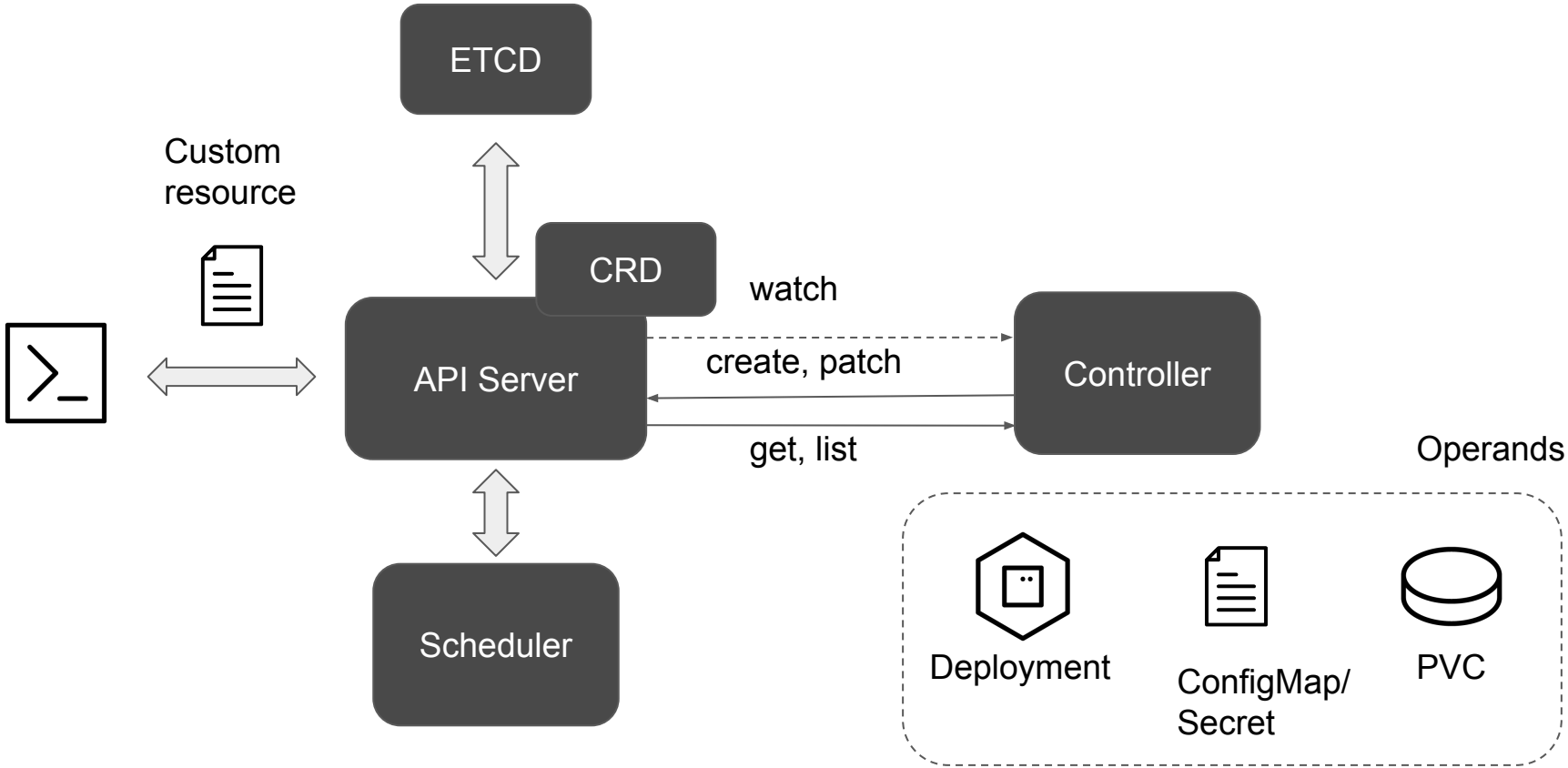
“

How to automate operating applications? The Operator pattern!

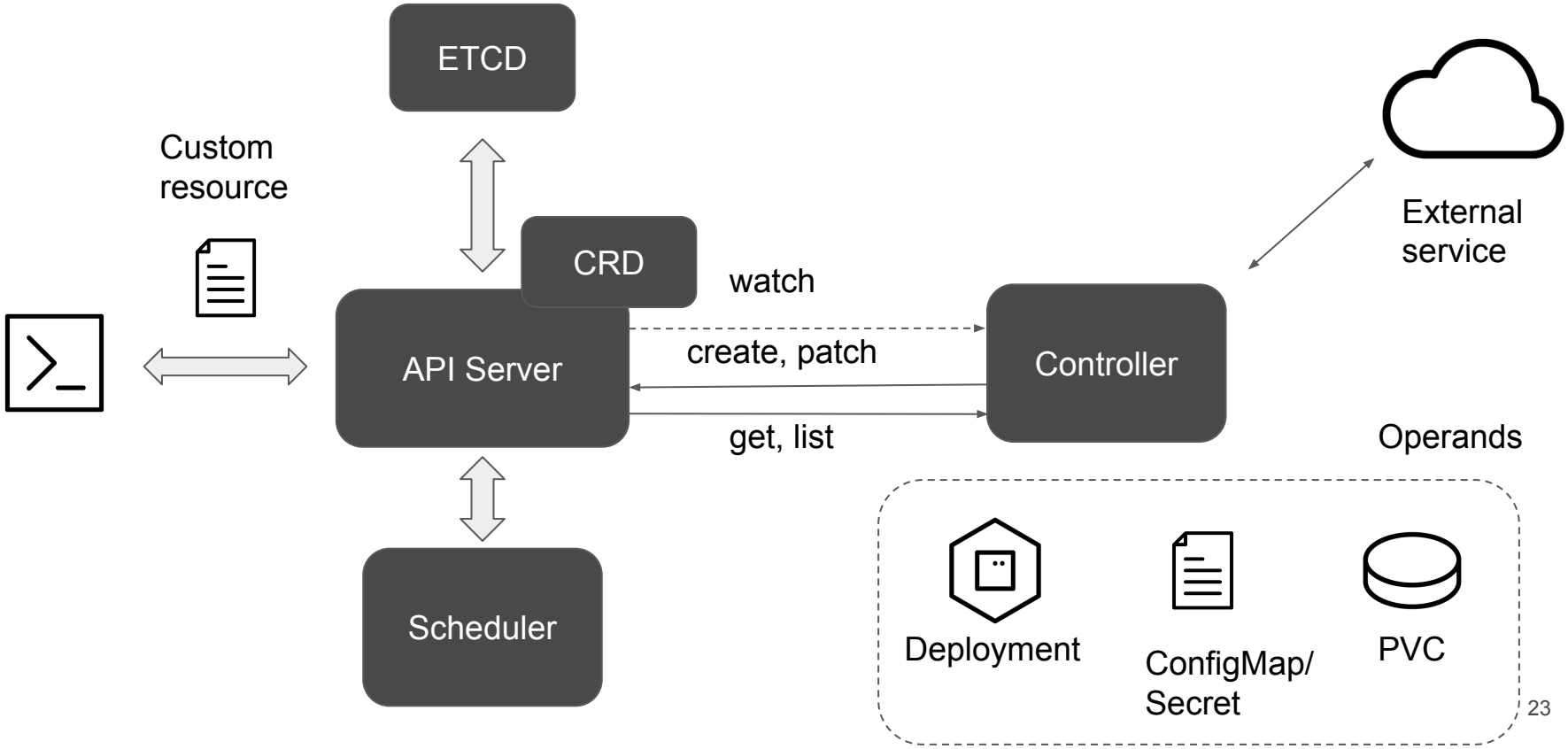
Operator

- It’s yet another containerized application!
- Has the knowledge of a specific business domain
- Manage the application lifecycle
- Leverage CRDs (Custom Resource Definition) to extend API server
- Takes care of one (or more) custom resources/objects
  - By having a controller for each resource
  - Creating native Kubernetes resources … aka “operands”
  - Leveraging built-in controllers via API server interaction

Operator


ETCD
Custom resource
CRD
watch
API Server
create, patch
Controller
get, list
Operands
Scheduler
Deployment
ConfigMap/ Secret
PVC

Operator


ETCD
Custom resource
CRD
watch
API Server
create, patch
Controller
External service
get, list
Operands
Scheduler
Deployment
ConfigMap/ Secret
PVC

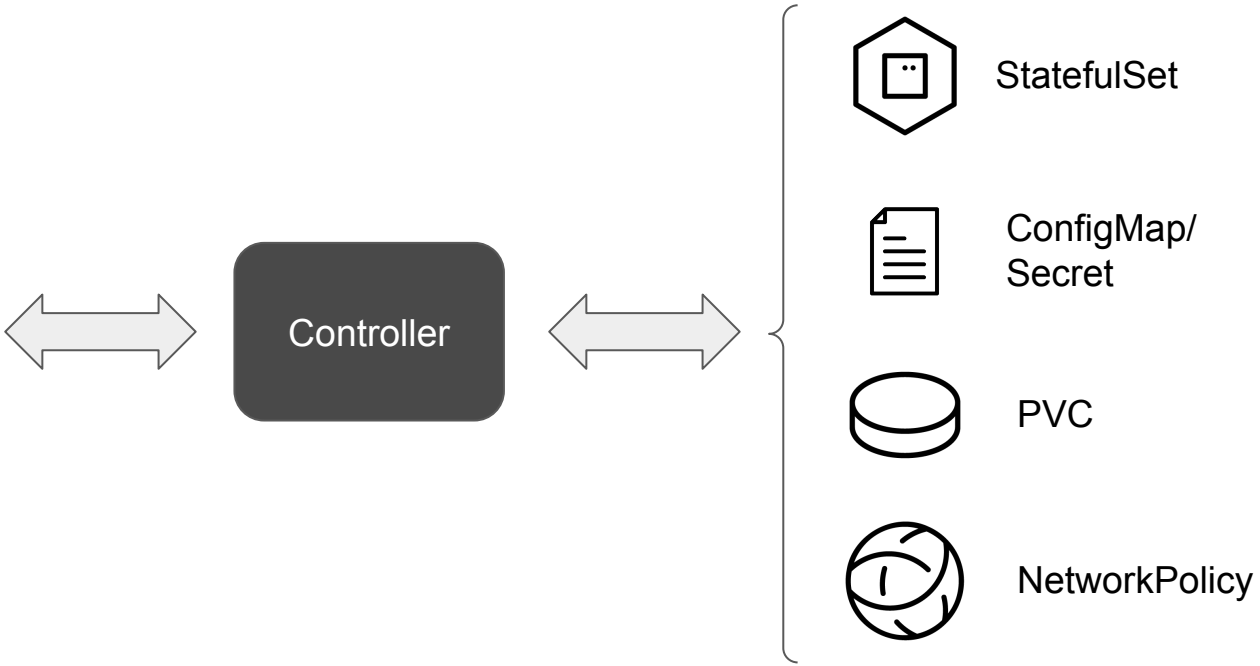
From the Custom Resource Definition ...

- It's a … Kubernetes resource!
- Declare a new Kubernetes "kind"
  - group
  - versions
- Define the new "kind" structure using an OpenAPI schema
  - spec
  - status

```yaml
apiVersion: apiextensions.k8s.io/v1
kind: CustomResourceDefinition
metadata:
  name: kafkas.kafka.strimzi.io
spec:
  group: kafka.strimzi.io
  names:
    kind: Kafka
    listKind: KafkaList
  ...
  versions:
  - name: v1beta2
    schema:
      openAPIV3Schema:
        type: object
        properties:
          spec:
            # spec definition with for
            # the custom resource
            ...
          status:
            # status definition reported back
            # in the custom resource
```

… to the Custom Resource

```yaml
apiVersion: kafka.strimzi.io/v1beta2
kind: Kafka
metadata:
  name: my-cluster
spec:
  kafka:
    version: 2.8.0
    replicas: 3
    listeners:
      - name: plain
        port: 9092
        type: internal
        tls: false
      - name: tls
        port: 9093
        type: internal
        tls: true
    config:
      log.message.format.version: "2.8"
      inter.broker.protocol.version: "2.8"
      ...
    storage:
      type: ephemeral
  zookeeper:
    replicas: 3
    storage:
      type: ephemeral
status:
  ...
  ...
```



Operator

- Each “internal” controller watches a corresponding custom resource
  - Create/update/delete the corresponding “operands” as native Kubernetes resources but … could be other custom resources for other operators :-)
- Watch the “operands”
  - They can be touched only by operator controllers … not humans
  - Reverts back any manual changes
  - “Owned by” the custom resource, leveraging Kubernetes garbage collection
- Can interact with external service
  - A custom resource could be related to handle a non-Kubernetes service (i.e. Azure Service Operator, …)

Operator

- Configuration
  - Simplifies configuration of a complex application in a single and meaningful place
- Installation and upgrade
  - Executes installation but mostly is able to run a potential complex upgrade procedure
- Security
  - Allows to take care of security related stuff like encryption, access control, network policies
- Scaling
  - Enable scaling and maybe, depending on application, autoscaling

Why? What about Helm Charts?

Helm Charts

- “Package manager” for Kubernetes
  - Charts = template + values
- Rely on Kubernetes built-in resources (i.e. Deployment, ConfigMap, …)
- Simplify to write YAMLs with parameters via templating
- Simplify day-1 operation, for deploying applications
- Key problems fixed
  - Deploy same application with different configuration
  - Deploy same application on different environments

Operators

- Control “life cycle” of Kubernetes workloads
  - Operator = CRD(s) + Controller(s)
- Extend the Kubernetes API with CRDs (Custom Resource Definitions)
- Simplify to write one (or a few) “custom resource” related YAMLs
- Acting since day-1 to day-2 operation from deployment to upgrades, through manage
- … deployable via an Helm Charts :-)
  - Helm guarantees CRDs are installed before operator
  - Operator configurable via values or ConfigMap

“

You convinced me! How to start?

Developing an operator

- Operator framework
  - Toolkit to manage Kubernetes operator
  - SDK for writing operators in Golang
  - Operator Lifecycle Manager (OLM) to handle … operators
  - Operator registry to provide operators to OLM
  - … and much more
  - https://sdk.operatorframework.io/
- Java Operator SDK
  - Writing operators in Java
  - Support for Quarkus
  - https://javaoperatorsdk.io/

OPERATOR
FRAMEWORK

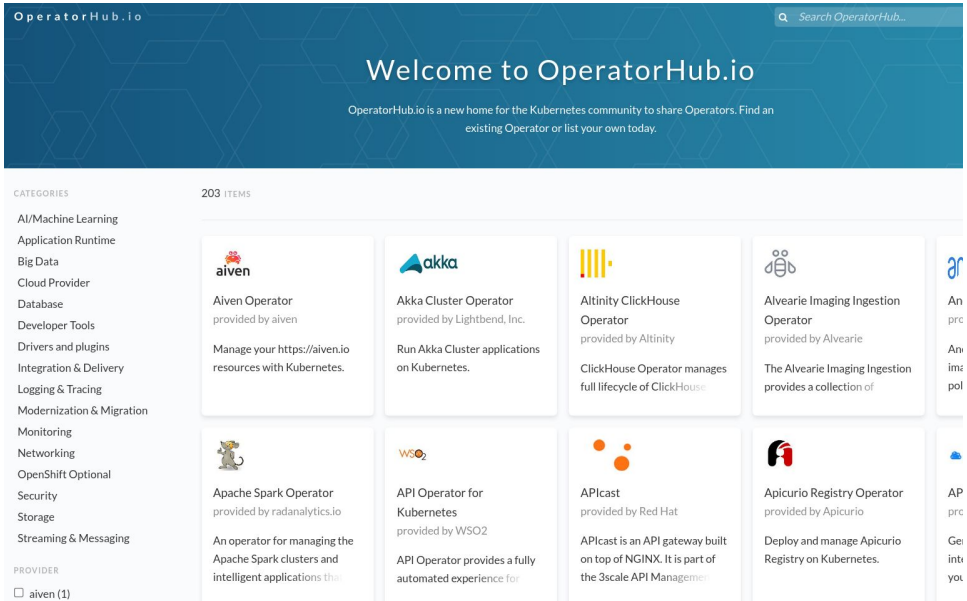
[JAVA OPERATOR SDK]

Operator Lifecycle Manager (OLM)

- An operator to rule them all!
  - An operator handling … operators!
- Define a packaging format
- Expose a catalog
- Handle operator updates
- Provided out of the box with OpenShift (enterprise Kubernetes distribution by Red Hat) but installable on vanilla Kubernetes

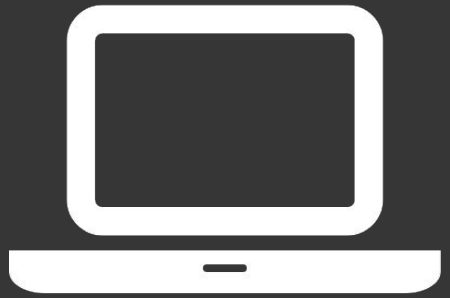
OperatorHub.io

- Home for Kubernetes operators
  - A lot of categories (Database, Streaming & messaging, Logging & Tracing, …)
- Installation via Helm Charts or YAML files
- You can develop your own and provide to the community
- https://operatorhub.io/



Strimzi: The Apache Kafka operatore

- Open source project licensed under Apache License 2.0
- Focuses on running Apache Kafka on Kubernetes
  - Container images for Apache Kafka, Apache ZooKeeper and other components
  - Operators for deploying, managing and configuring Kafka clusters
- Provides a Kubernetes-native experience
  - Not only Kafka clusters, but also users, topics and the rest of Kafka ecosystem
- CNCF sandbox project since September 2019
- http://strimzi.io



Demo time

Thank you

Any questions?


ppatierno


@ppatierno


https://www.linkedin.com/in/paolopatierno/