



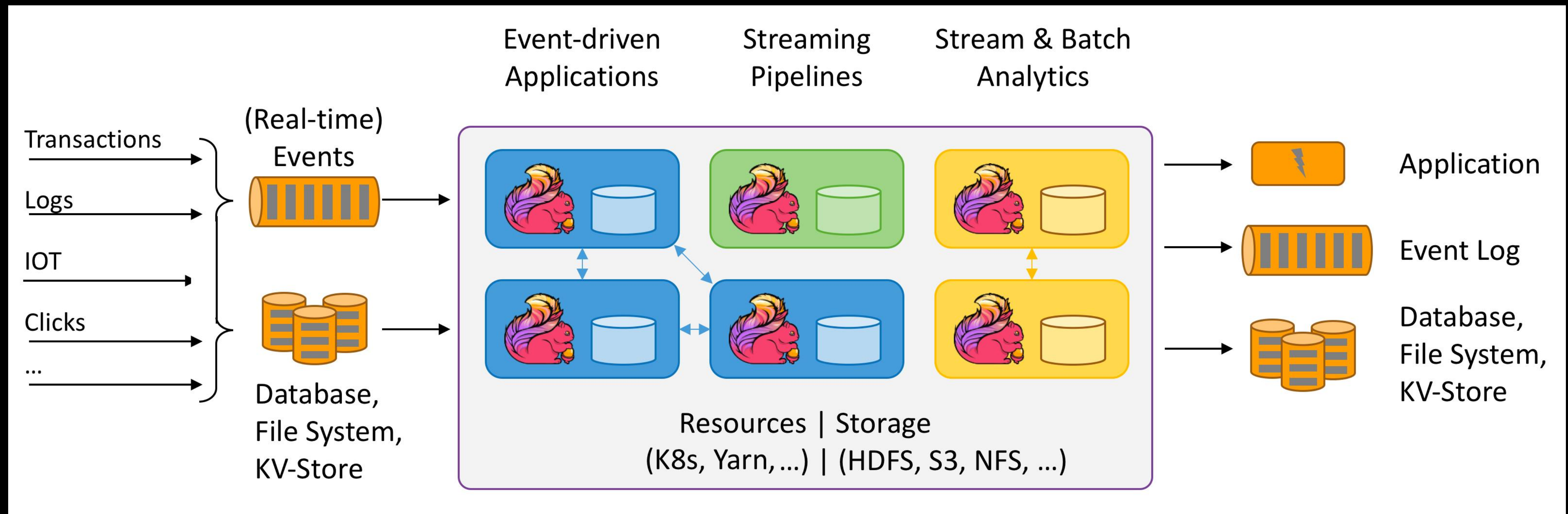
# Running Apache Flink in Any Cloud

Marton Balassi

Software Engineer, Apple | Apache Flink PMC member

# Flink in a nutshell

A distributed processing engine for stateful computations



# Cloud agnostic requirements

- Orchestration layer
- Scheduler
- Logging
- Metrics
- Local dev environment

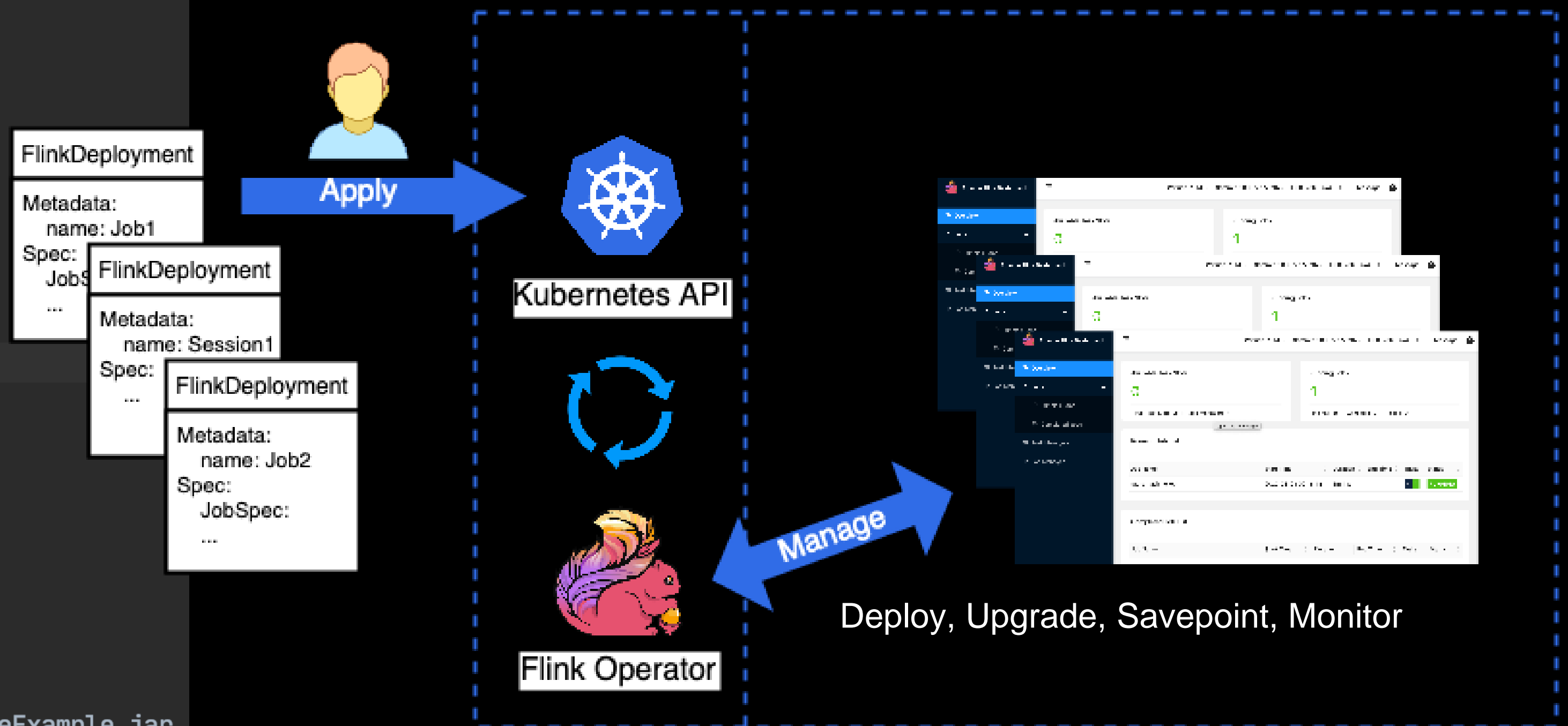


# Flink Kubernetes Operator

# Flink Kubernetes Operator

Manage Flink Application and Session clusters/jobs

```
apiVersion: flink.apache.org/v1beta1
kind: FlinkDeployment
metadata:
  name: basic-example
spec:
  image: flink:1.15
  flinkVersion: v1_15
  flinkConfiguration:
    taskmanager.numberOfTaskSlots: "2"
  serviceAccount: flink
  jobManager:
    resource:
      memory: "2048m"
      cpu: 1
  taskManager:
    resource:
      memory: "2048m"
      cpu: 1
  job:
    jarURI: local:///opt/flink/examples/streaming/StateMachineExample.jar
    parallelism: 2
    upgradeMode: stateless
```



# Why another Flink operator

Existing OSS solutions

- Fragmentation, unreliable maintenance

- Contribution barrier

Apache Flink needs a community supported operator

- Stable and open core logic

- Platforms should differentiate on a higher level

# Project Status


4 open source releases

~400 commits from 34 contributors


In production for months already...


Great initial adoption



Alibaba, AWS, IBM, VMware and many more...

 [apache / flink-kubernetes-operator](#) Public

Apache Flink Kubernetes Operator

 [flink.apache.org/](https://flink.apache.org/)

 Apache-2.0 license

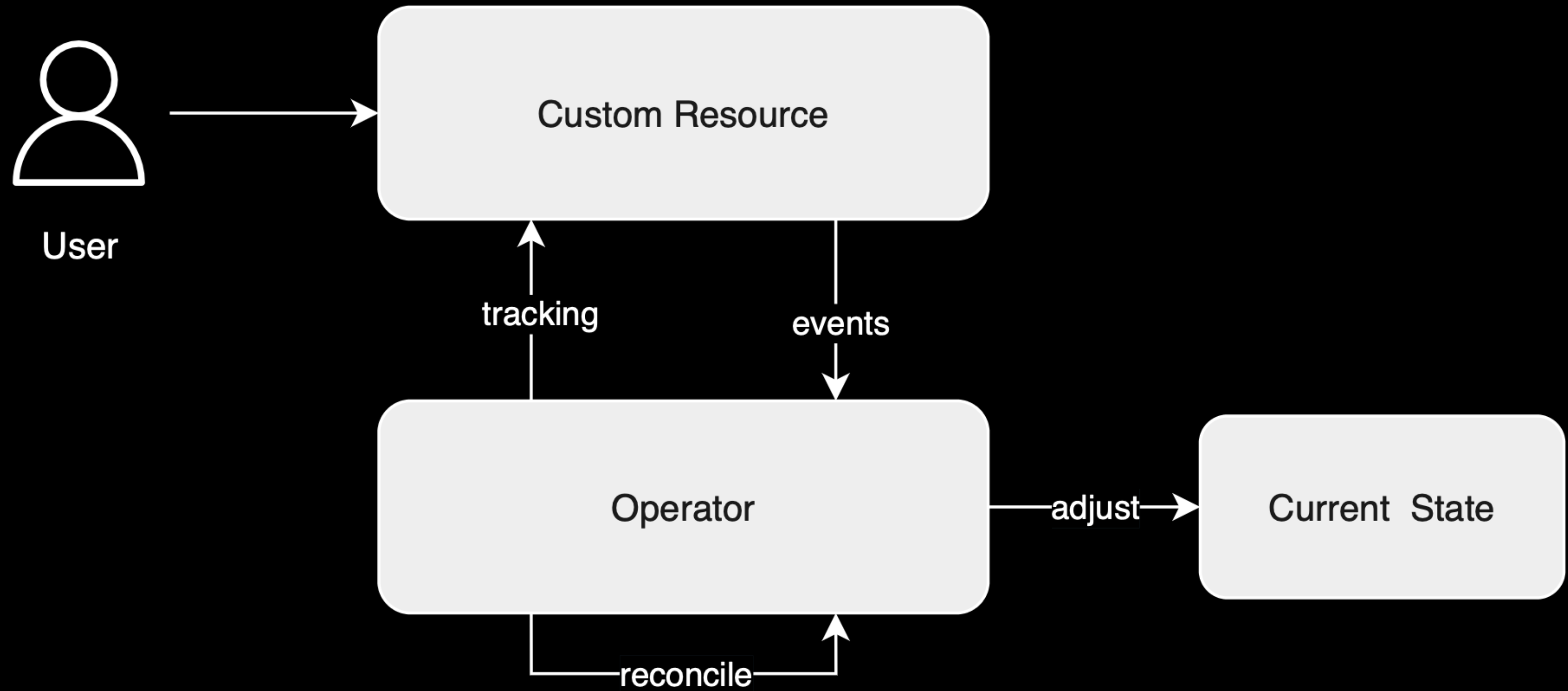
 284 stars  135 forks



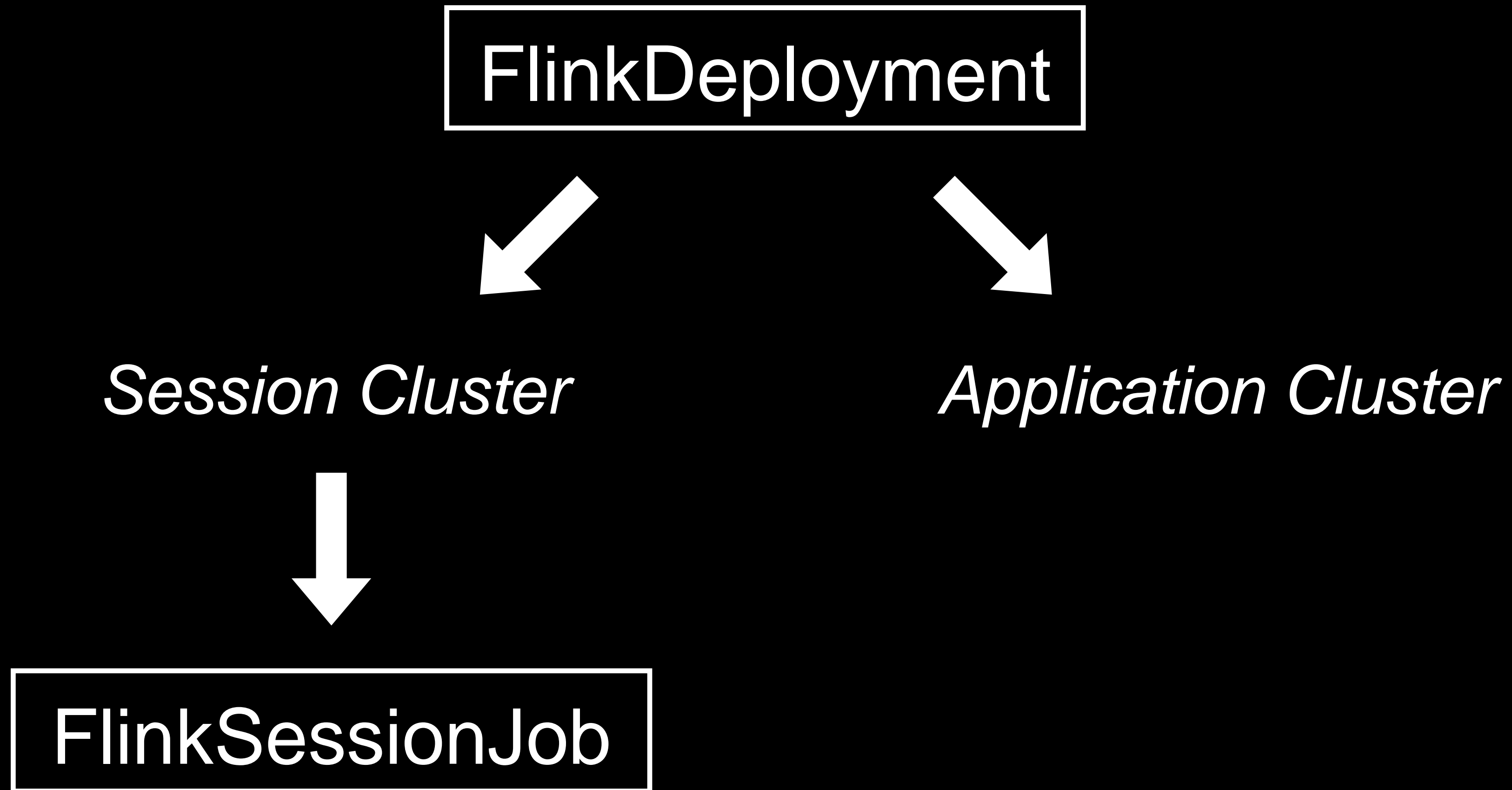
Demo

# Kubernetes Operator Design

# Control Loop



# Supported Deployment Types



# Kubernetes Operator Features

# Lifecycle Management

Running, Suspending and Deleting Applications

Stateful Application Upgrades

Savepoint Management

# Metrics & Logging

Integration with Flink metric system and reporters

Detailed operator metrics

- Resource lifecycle metrics

- Kubernetes client metrics

- JVM metrics

- JOSDK metrics

Detailed Resource and Operator level logging with MDC

# Kubernetes Events

```
replicas: 1
Events:
  Type     Reason          Age    From                    Message
  ----     -
  Normal   Submit          12m    JobManagerDeployment    Starting deployment
  Normal   StatusChanged   7m12s  Job                     Job status changed from RECONCILING to CREATED
  Normal   StatusChanged   6m57s  Job                     Job status changed from CREATED to RUNNING
```

```
kubectl describe flinkdeployment basic-example
```

Events can be forwarded to infrastructure specific collectors



# Robust & Flexible Architecture

Zero downtime config changes

- Dynamic Operator config updates

- Dynamic watched namespaces

Stateless architecture

- Operator only depends on CR status

Label selectors for multiple environments

# Summary

Kubernetes is a battle-tested, commonly supported container orchestrator platform across clouds (and on-premise).

The Flink community offers a Flink Kubernetes Operator to manage your Flink applications' lifecycle.

Production-readiness and productivity requirements can be implemented in a cloud agnostic way via existing Flink and Kubernetes tooling.

Q&A