The Institute
for Quantitative Social Science
at Harvard University

What's New with Apache Spark on Kubernetes

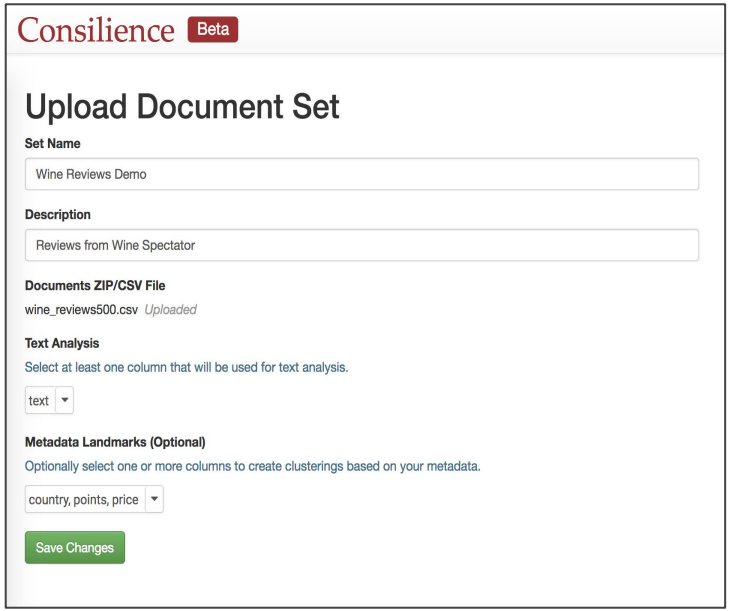
Ellen Kraffmiller - Senior Software Engineer
Robert Treacy - Senior Software Architect

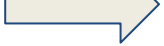
The Institute for Quantitative Social Science at Harvard University

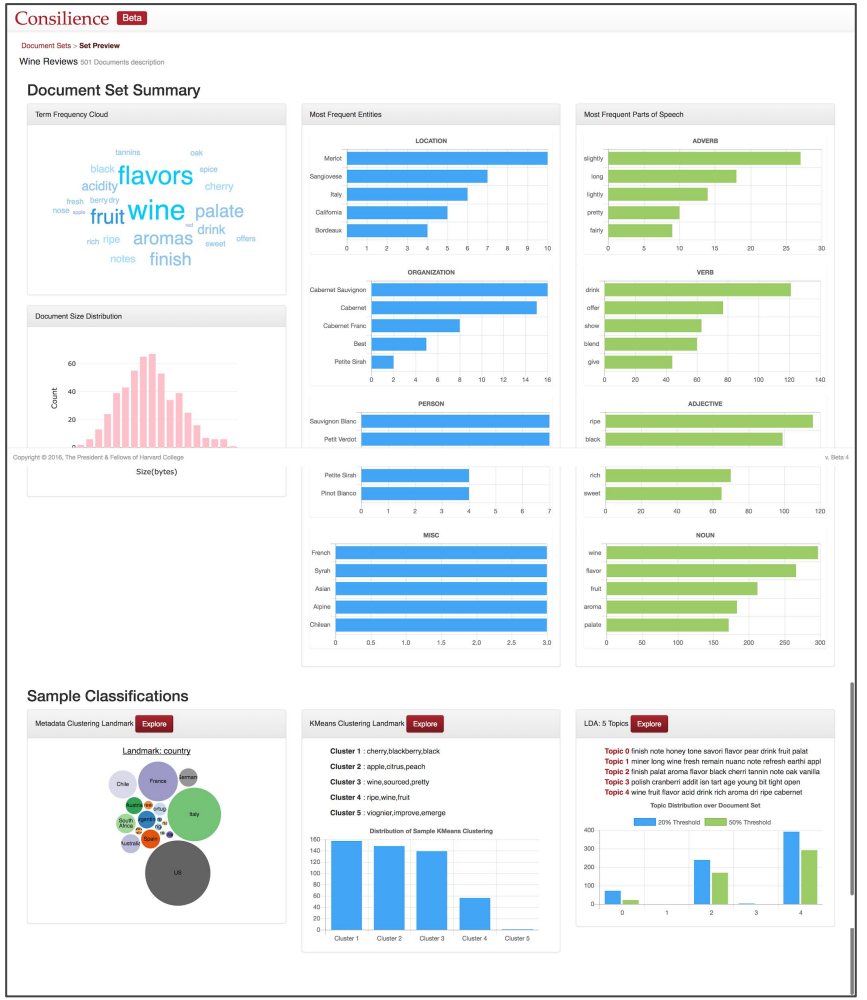
Agenda

- Who we are/Why we are interested in Spark on Kubernetes
- Overview of Spark on Kubernetes implementation
- What’s new since last year (Spark 2.4 features)
- What’s coming in Spark 3.0
- Spark Operator for Kubernetes
  - For configuring your Spark app
  - For monitoring your Spark app
- Demo of Spark Operator running Spark ML application

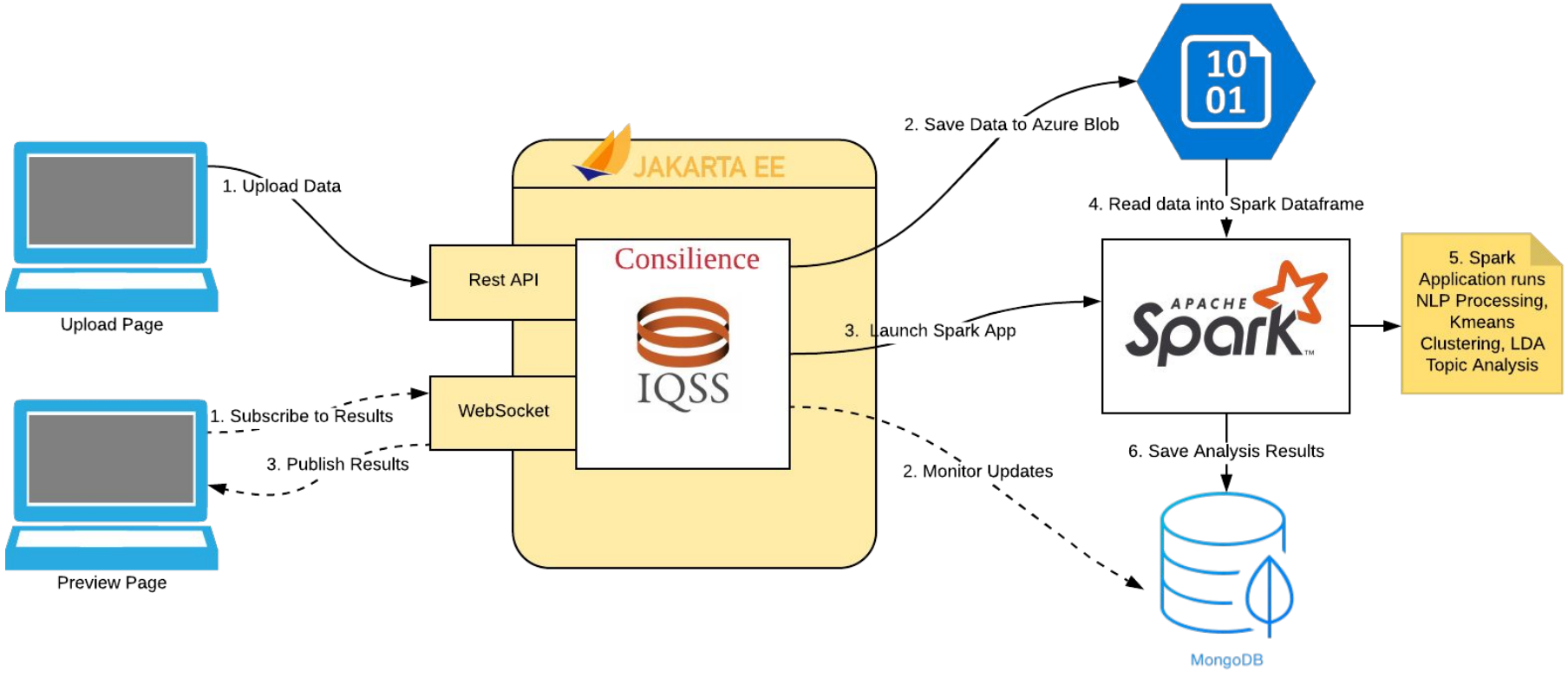
Consilience: Text Analysis Tool for Researchers







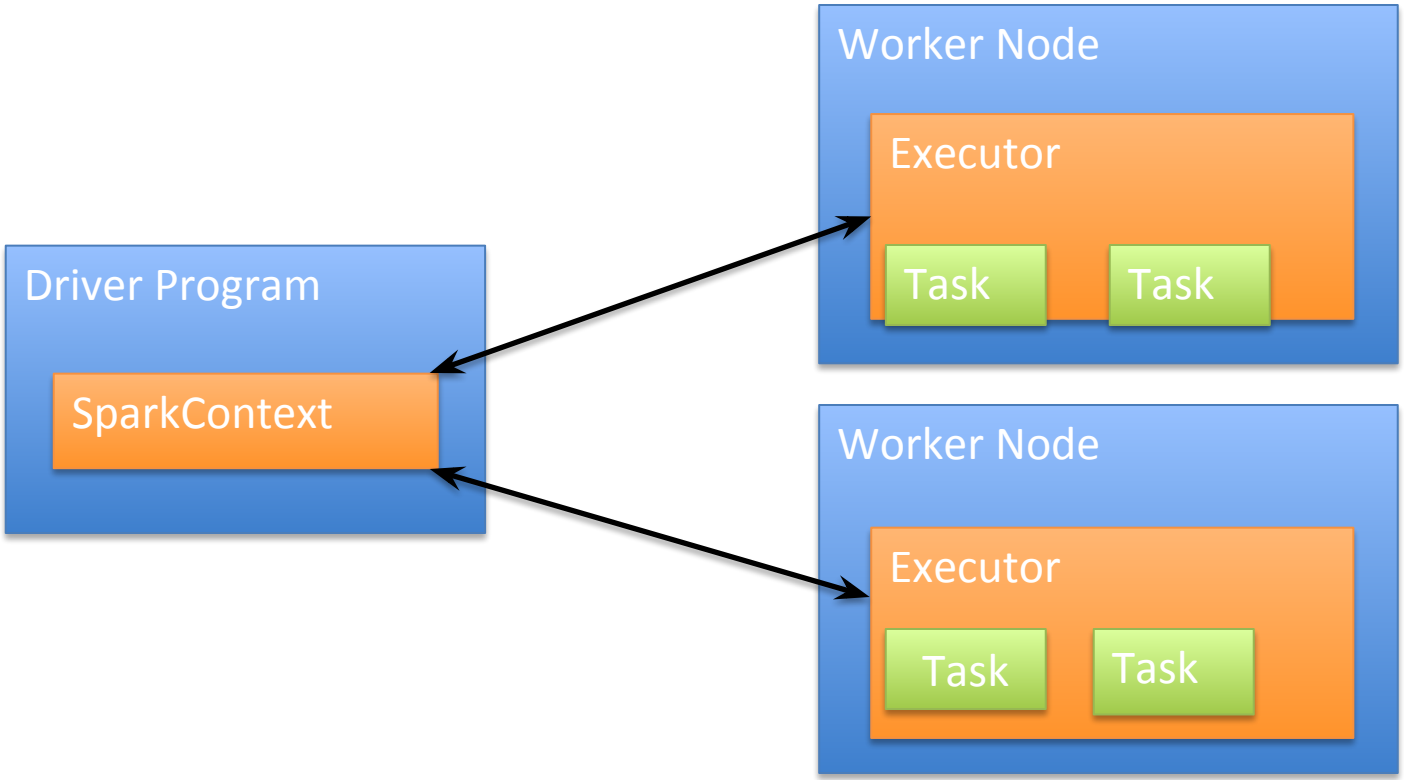
Consilience: Data Processing Workflow



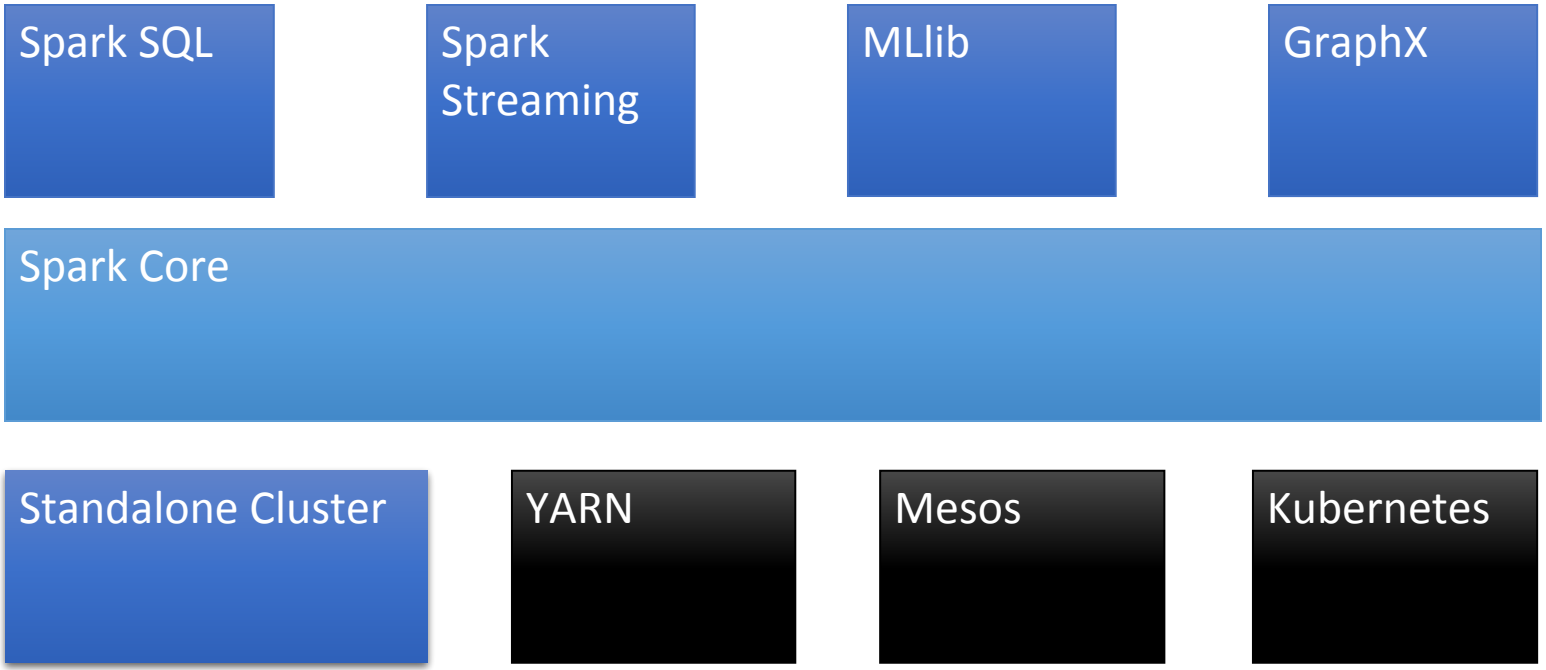
Apache Spark

- A fast and general engine for large scale data processing
- k8s support began in v 2.3.x
- 100x faster than Hadoop MapReduce in memory, 10x faster in disk

Cluster execution



Spark **Components**



Kubernetes

- Container-orchestration
- Operating system for clusters
  - Simplifies application deployment
  - Improves utilization of resources
- Standard for running distributed applications
- Containers

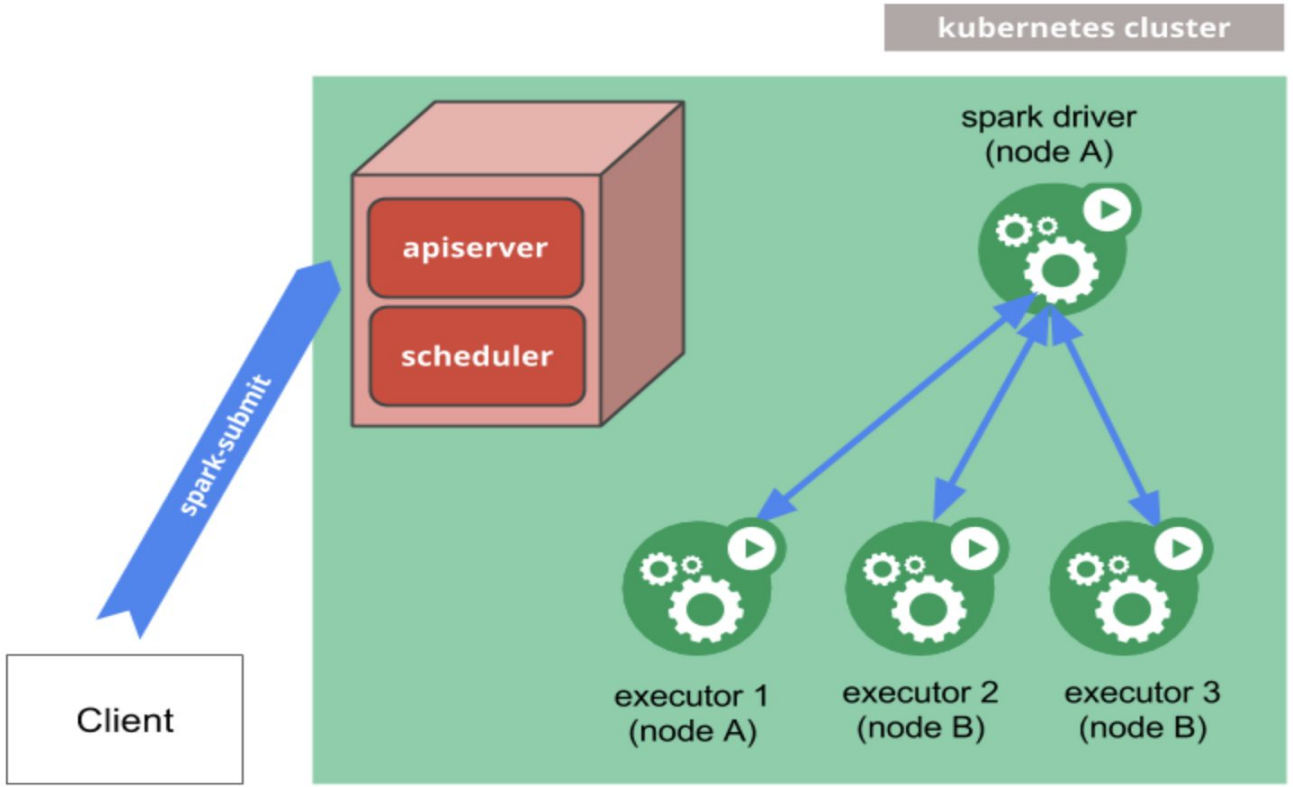
Kubernetes Automatic Scaling

- Google Kubernetes Engine
- Google Compute Engine
- Amazon Web Services
- Azure

Spark Kubernetes

- Spark v 2.3+
- Kubernetes 1.6+
- Docker

Spark on Kubernetes



tml

Reference:https://spark.apache.org/docs/latest/img/k8s-cluster-mode.png

Deploy to Kubernetes with spark-submit

```
bin/spark-submit --master
k8s://https://iqcodeone-73142460.hcp.eastus.azmk8s.io:443 --deploy-mode
cluster --name spark-pi --class org.apache.spark.examples.SparkPi --conf
spark.executor.instances=5 --conf
spark.kubernetes.container.image=bobtreacy/text-analysis --conf
spark.kubernetes.authenticate.driver.serviceAccountName=spark
local:///opt/spark/examples/jars/spark-examples_2.11-2.3.2.jar 50000
```

Spark 2.4

- PySpark bindings for K8S
- R bindings for K8S
- Support client mode for Kubernetes cluster backend
- Support for mounting K8S volumes

Future

Dynamic resource allocation and external shuffle service

Local file dependency management

Spark application management

Job queues and resource management

Spark 3.0

Pod Templates

GPU-aware scheduling

Sub path mounting

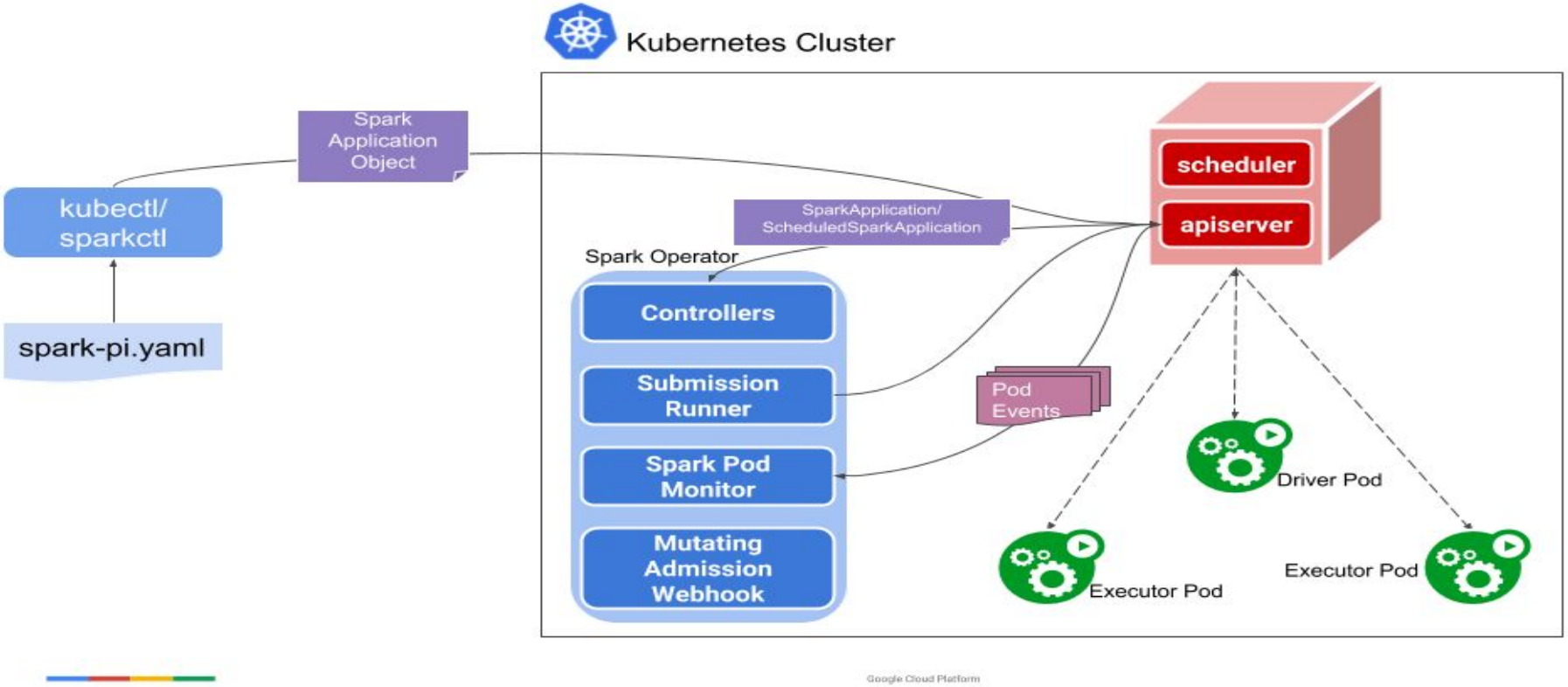
Improve behavior with dynamic allocation

Integration testing improvements

Kubernetes Operator for Apache Spark

- Google
  - Google Kubernetes Engine
  - Run on any Kubernetes 1.8+
- Supports Spark 2.3 +
- Customization of Spark pods
  - Kubernetes Mutating Admission Webhook

Kubernetes Operator for Spark



Reference: https://github.com/GoogleCloudPlatform/spark-on-k8s-operator/blob/master/docs/architecture-diagram.png

Spark & RBAC

If RBAC is enabled, the Spark driver pod needs permission to create and watch Spark executor pods:

```
$ kubectl create serviceaccount spark
```

```
$ kubectl create clusterrolebinding spark-role --clusterrole=edit
--serviceaccount=default:spark --namespace=default
```

Custom Resource Definitions (CRD)

- SparkApplication
  - kubectl apply -f <YAML file> or sparkctl create <YAML file>
- ScheduledSparkApplication
  - cron schedule

SparkApplication Spec

```
apiVersion: sparkoperator.k8s.io/v1beta1
kind: SparkApplication
metadata:
  name: spark-pi
  namespace: default
spec:
  type: Scala
  mode: cluster
  image: gcr.io/spark/spark:v2.4.0
  mainClass: org.apache.spark.examples.SparkPi
  mainApplicationFile: local:///opt/spark/examples/jars/spark-examples_2.11-2.4.0.jar
```

Mutating Admission Webhook

Optional - customize Spark driver and executor pods

Kubernetes 1.9+

kubectl apply -f manifest/spark-operator-with-webhook.yaml

--set enableWebhook=true (as a param for helm/install)

Mutating Admission Webhook use

- Mounting a ConfigMap
  - Spark Configuration Files
  - Hadoop Configuration Files
- Mounting Volumes
- Pod Security Context
- DNS Settings

Tolerations and Pod Affinity

Tolerations

- Node resources reserved for specific pods

Pod Affinity

- Grouping pods on specific node to work together

sparkctl

- Command line tool – like kubectl extended
  - Create, list, check status, get logs ...
- Flags
  - --namespace
  - --kubeconfig
- Mount local Hadoop config as a ConfigMap

Demo

https://github.com/ekraffmiller/CodeOne2019

Related Presentations

DEV2410 - Getting Started with Spark
DEV5377 - Our Experience Writing a Kubernetes Operator
DEV1641-Data Pipelines with Apache Spark

Thank you!

- Demo on Github:
  https://github.com/ekraffmiller/CodeOne2019

- Questions?