Airflow / Kubernetes
Running on and using k8s
Jed Cunningham
Staff Software Engineer @ astronomer.io

Committer / PMC
On Deck

**Airflow**
Operator / Executor?

**Kubernetes**
Tasks
Airflow itself

**Helm**
Chart options
KubernetesPodOperator
KubernetesExecutor
Which? Both?!
Operator

What does this task need to do?

Think function.

Executor

How do I run tasks?

Think environment.
Operator
What do I need to do?

Generic:
PythonOperator
BashOperator

AWS S3 == S3*Operator
SQL == PostgresOperator/etc
Hive == HiveOperator
Email == EmailOperator
...

`docker run ...` == KPO
E.g. run something in a container
Generally NOT Airflow things!
Executor
How do I do it?

Pick 1: Speed or Isolation
Any combo works!

Don’t force associations.
Kubernetes Executor

Scheduler

Persistent

Ephemeral

Worker (KPO)

hello-world
HOW?

You could...

● Write your own k8s manifest
● Write your own helm chart
● Use an existing chart
Which?

(At least) 3 chart options:

- Bitnami
- User-community (old helm stable)
- Official
Official Chart

- By the community, for the community
- Over 90 contributors!
- Production focused
- 2.0+ (and 1.10+)
- Supports (nearly) all executors
- To many features to list...
helm repo add airflow https://airflow.apache.org
"airflow" has been added to your repositories
helm upgrade --install airflow airflow/airflow
Release "airflow" does not exist. Installing it now.
NAME: airflow
LAST DEPLOYED: Mon May 23 17:00:33 2022
NAMESPACE: default
STATUS: deployed
REVISION: 1
TEST SUITE: None
NOTES:
Thank you for installing Apache Airflow 2.3.0!
```
~ kubectl get pods

<table>
<thead>
<tr>
<th>NAME</th>
<th>READY</th>
<th>STATUS</th>
<th>RESTARTS</th>
<th>AGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>airflow-postgresql-0</td>
<td>1/1</td>
<td>Running</td>
<td>0</td>
<td>43s</td>
</tr>
<tr>
<td>airflow-redis-0</td>
<td>1/1</td>
<td>Running</td>
<td>0</td>
<td>43s</td>
</tr>
<tr>
<td>airflow-scheduler-5db94dd79c-hbd8h</td>
<td>2/2</td>
<td>Running</td>
<td>0</td>
<td>43s</td>
</tr>
<tr>
<td>airflow-statsd-5bcb9dd76-xz4wb</td>
<td>1/1</td>
<td>Running</td>
<td>0</td>
<td>43s</td>
</tr>
<tr>
<td>airflow-triggerer-6cf7cfd5f7-gzjxm</td>
<td>1/1</td>
<td>Running</td>
<td>0</td>
<td>43s</td>
</tr>
<tr>
<td>airflow-webserver-9f89969f-5xg8b</td>
<td>1/1</td>
<td>Running</td>
<td>0</td>
<td>43s</td>
</tr>
<tr>
<td>airflow-worker-0</td>
<td>2/2</td>
<td>Running</td>
<td>0</td>
<td>43s</td>
</tr>
</tbody>
</table>
```
KPO
`pod_template_file`
`pod_mutation_hook`

KubernetesExecutor
`executor_config.pod_override`
Thanks!

Q/A at:

#airflow-kubernetes
#helm-chart-official