

Airflow / Kubernetes

Running on and using k8s



Jed Cunningham

Staff Software Engineer @ astronomer.io

Committer / PMC



ASTRONOMER

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

	LocalExecutor	CeleryExecutor	KubernetesExecutor
Fast	✓	✓	
Simple	✓✓	✓	
Isolation			✓
Custom Resources			✓

Any combo works!

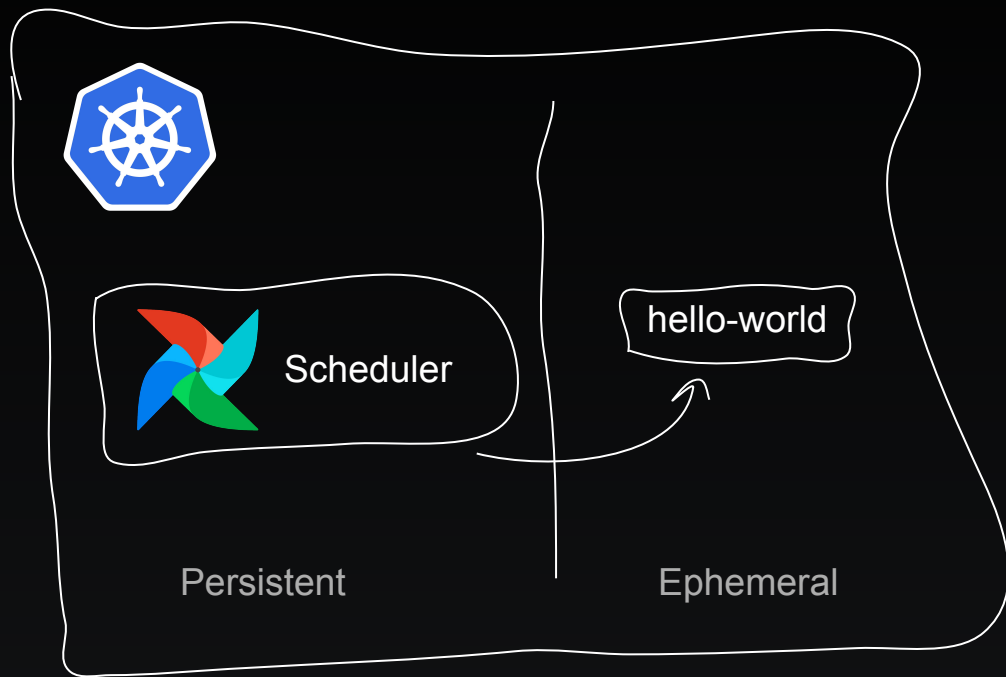
Don't force associations.

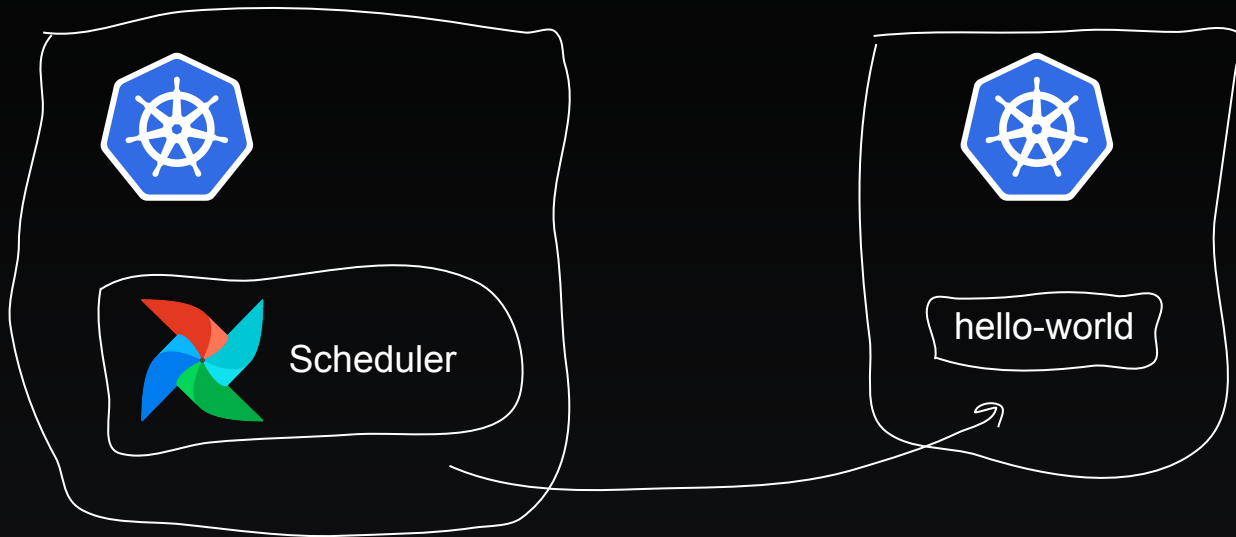


LocalExecutor

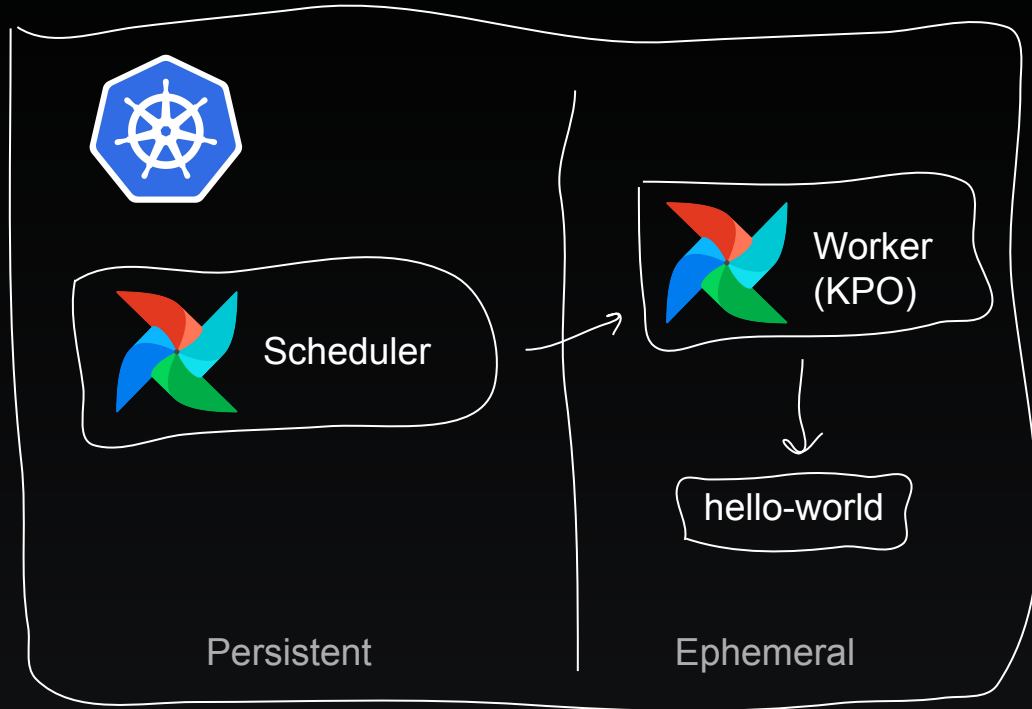
```
KubernetesPodOperator(  
    namespace="default",  
    image="hello-world",  
    name="hello-world",  
    task_id="hello-world",  
)
```

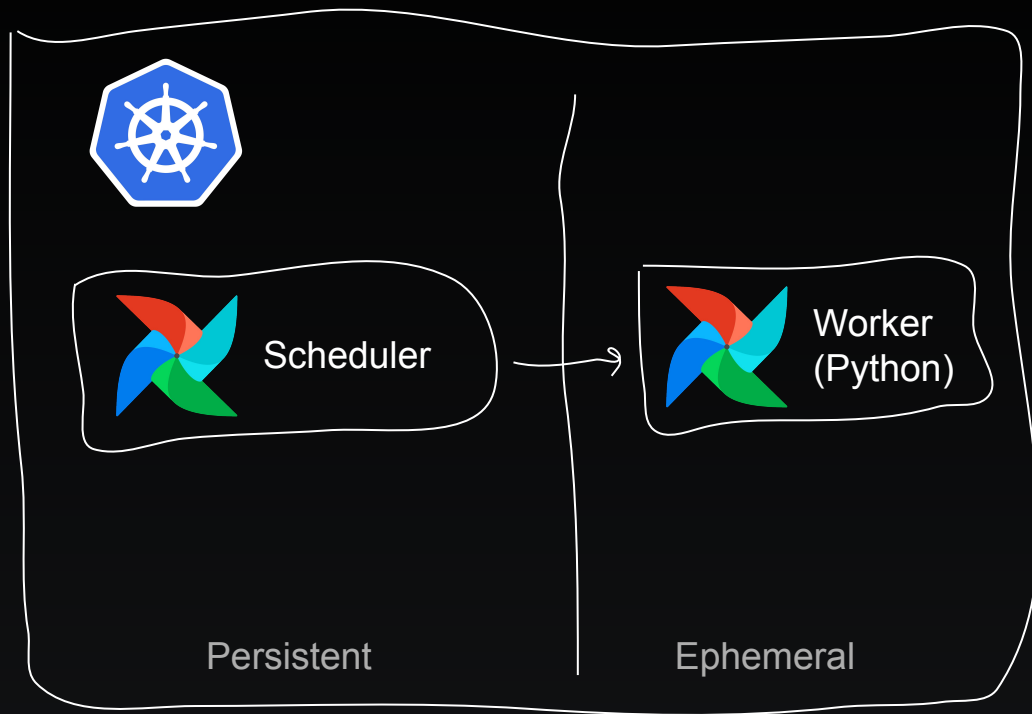






KubernetesExecutor





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
```

NAME	READY	STATUS	RESTARTS	AGE
airflow-postgresql-0	1/1	Running	0	43s
airflow-redis-0	1/1	Running	0	43s
airflow-scheduler-5db94dd79c-hbd8h	2/2	Running	0	43s
airflow-statsd-5bcb9dd76-xz4wb	1/1	Running	0	43s
airflow-triggerer-6cf7cfd5f7-gzjxm	1/1	Running	0	43s
airflow-webserver-9f89969f-5xg8b	1/1	Running	0	43s
airflow-worker-0	2/2	Running	0	43s

KPO

KubernetesExecutor

``pod_template_file``

``executor_config.pod_override``

``pod_mutation_hook``

Thanks!

Q/A at:

#airflow-kubernetes
#helm-chart-official