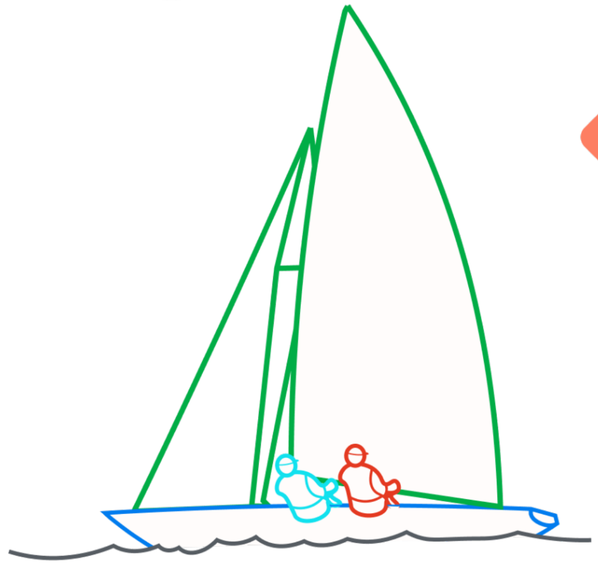
Better Support for Using Multiple Namespaces with KubernetesExecutor



Xiaodong Deng (@XD-DENG)

Software Engineer at Apple
Apache Airflow PMC member & Committer

Airflow Summit
Let's flow together

September 19-21, 2023,
Toronto, Canada

About myself

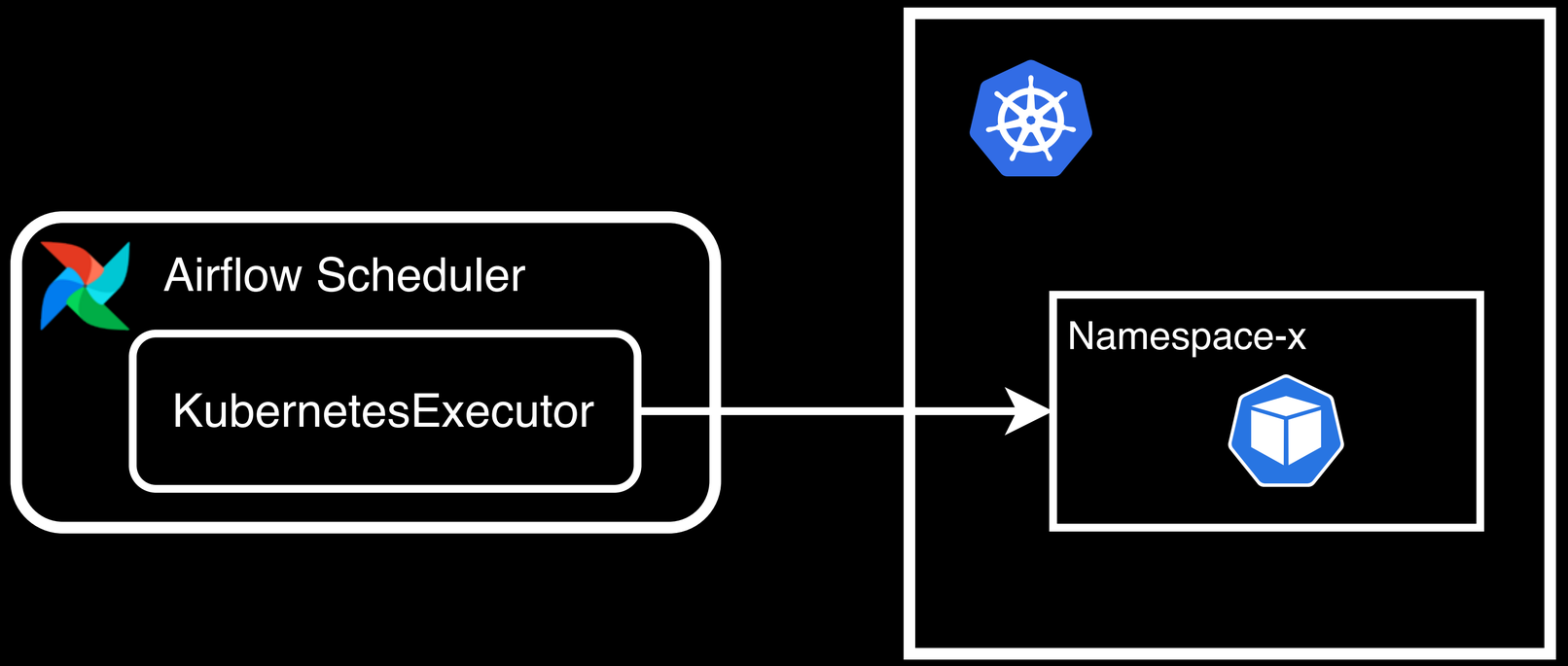
- Software Engineer at Apple
- Started to work on Airflow since 2018
- Airflow Committer since March 2019
- Airflow PMC since December 2020

About myself

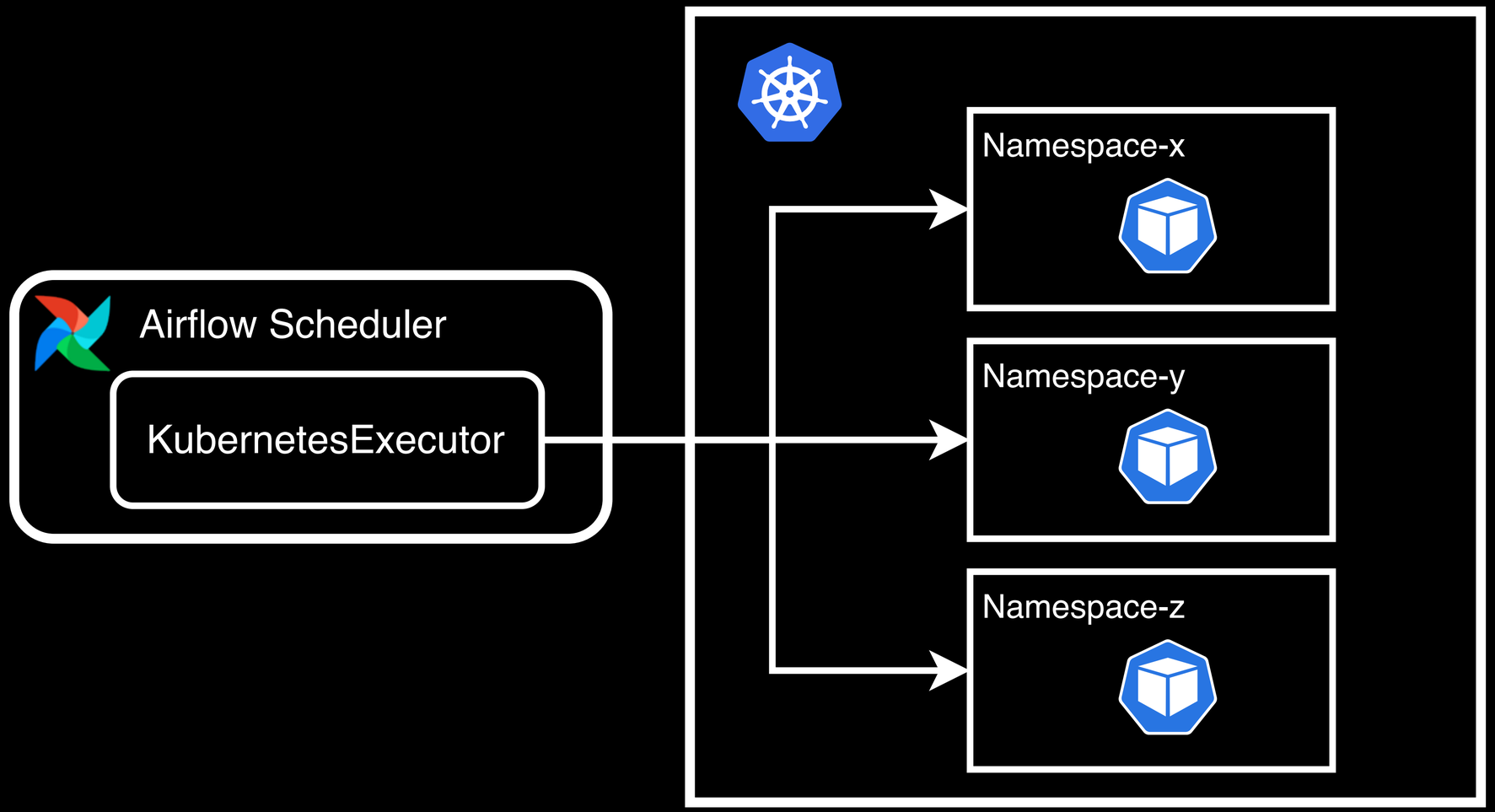
Other than **Apache Airflow**,

I also deal with **Airflow** *physically*.

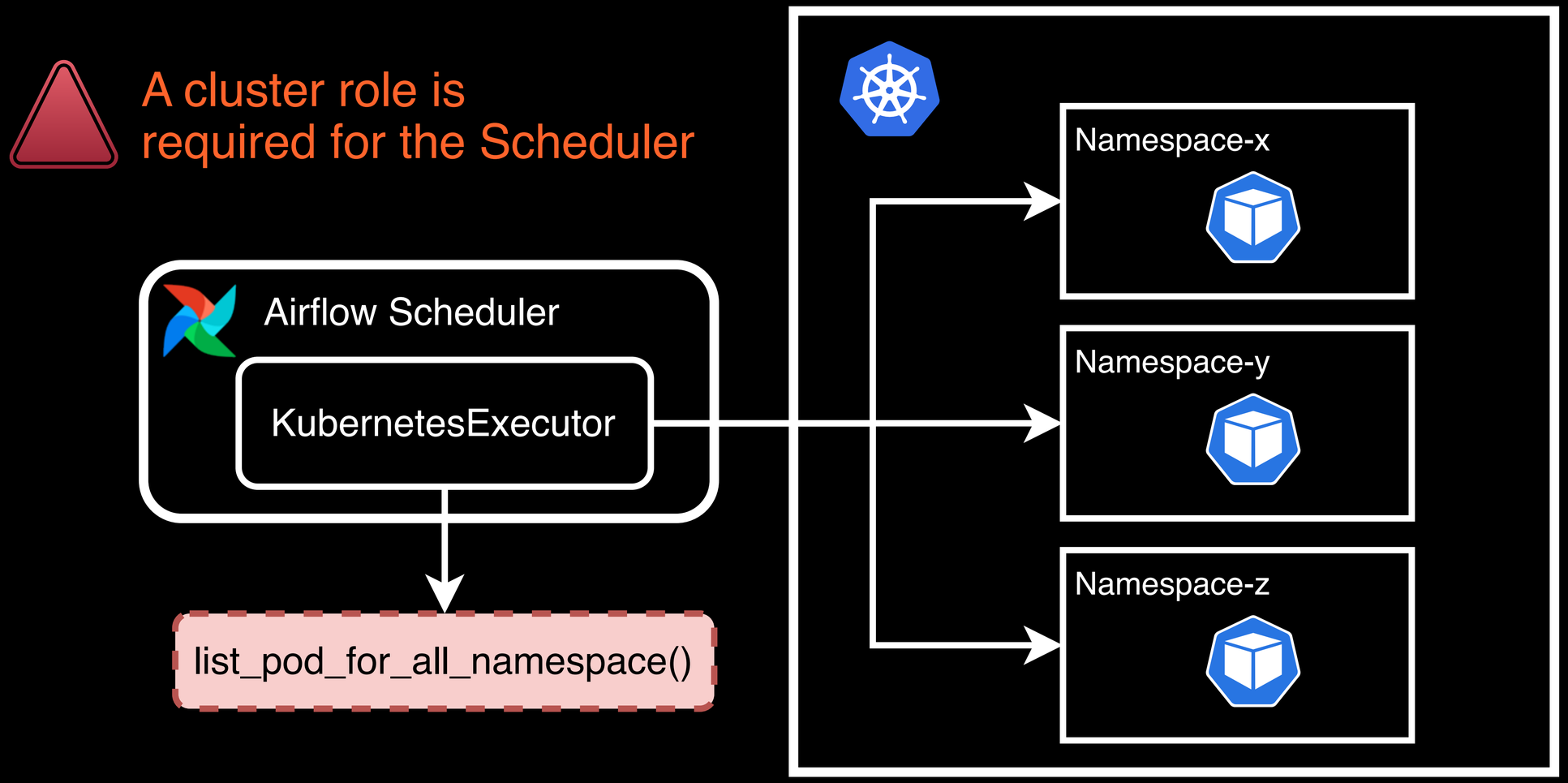
KubernetesExecutor & How was it like?



KubernetesExecutor & How was it like?



KubernetesExecutor & How was it like?



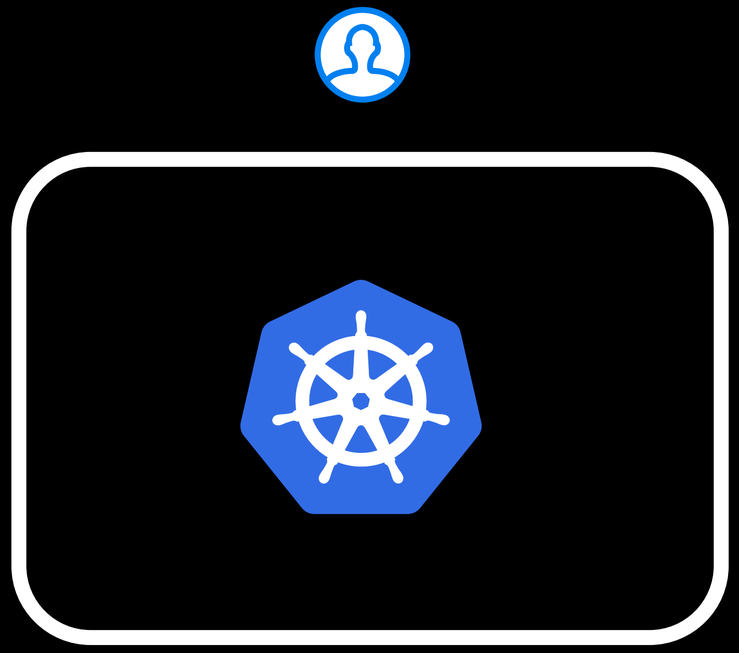
KubernetesExecutor & How was it like?

Airflow 2.5.3 & lower

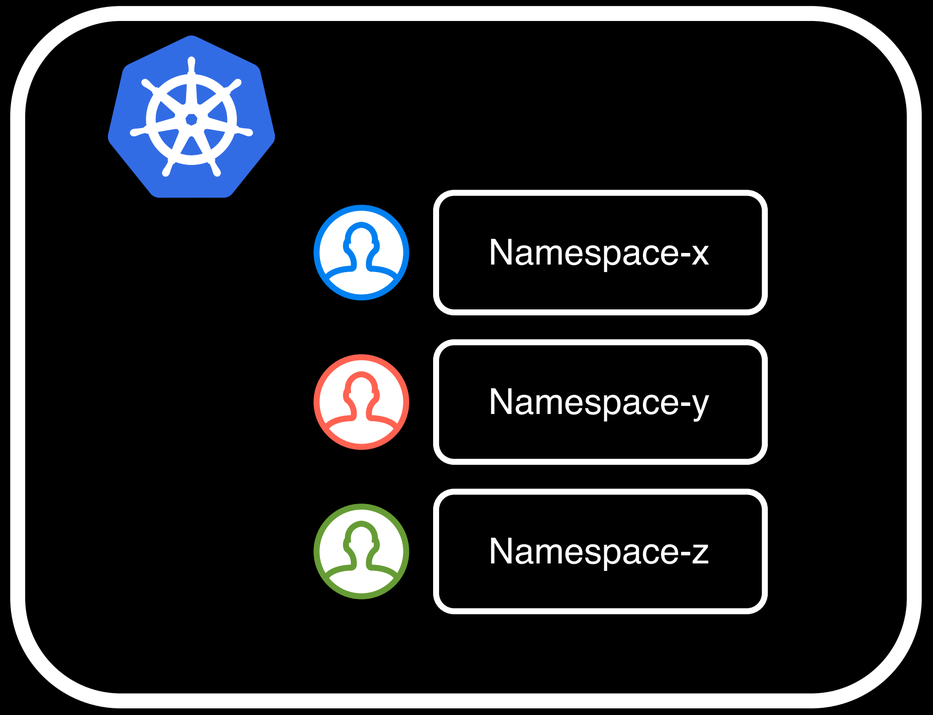
```
# Allows users to launch pods in multiple namespaces.
# Will require creating a cluster-role for the scheduler
multi_namespace_mode = False
```

What was the pain/Why should there be a change?

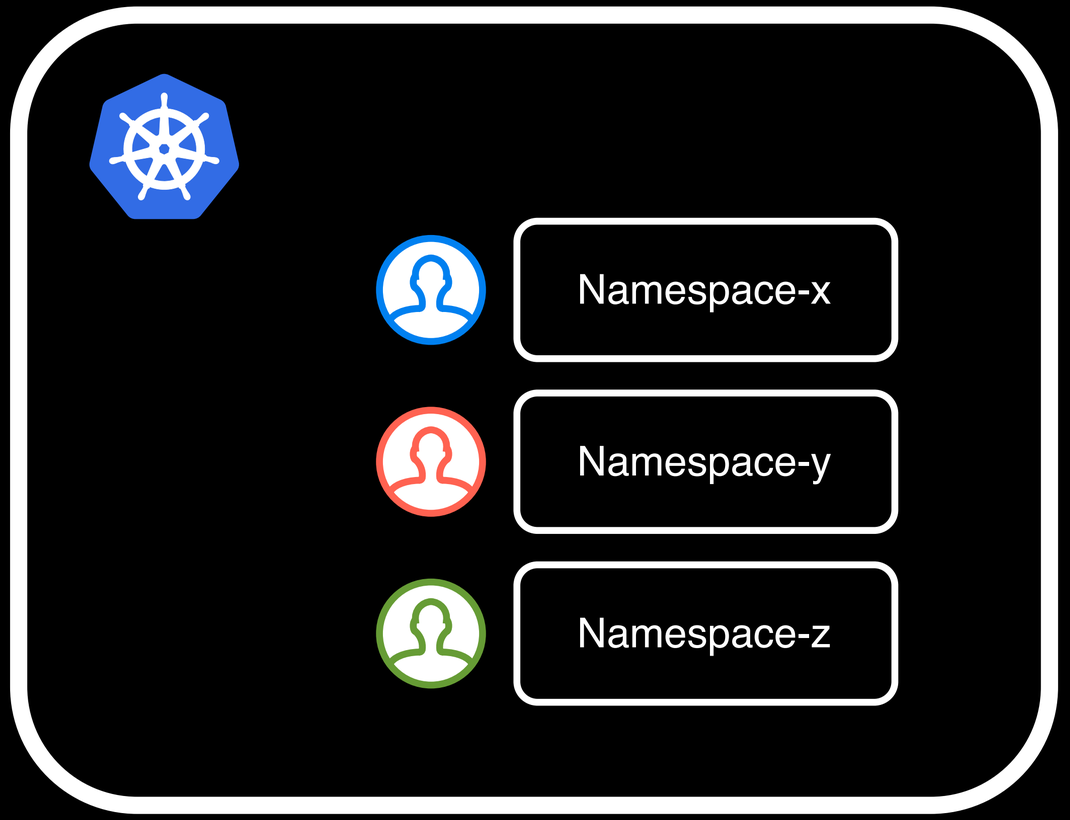
What you normally see in a demo



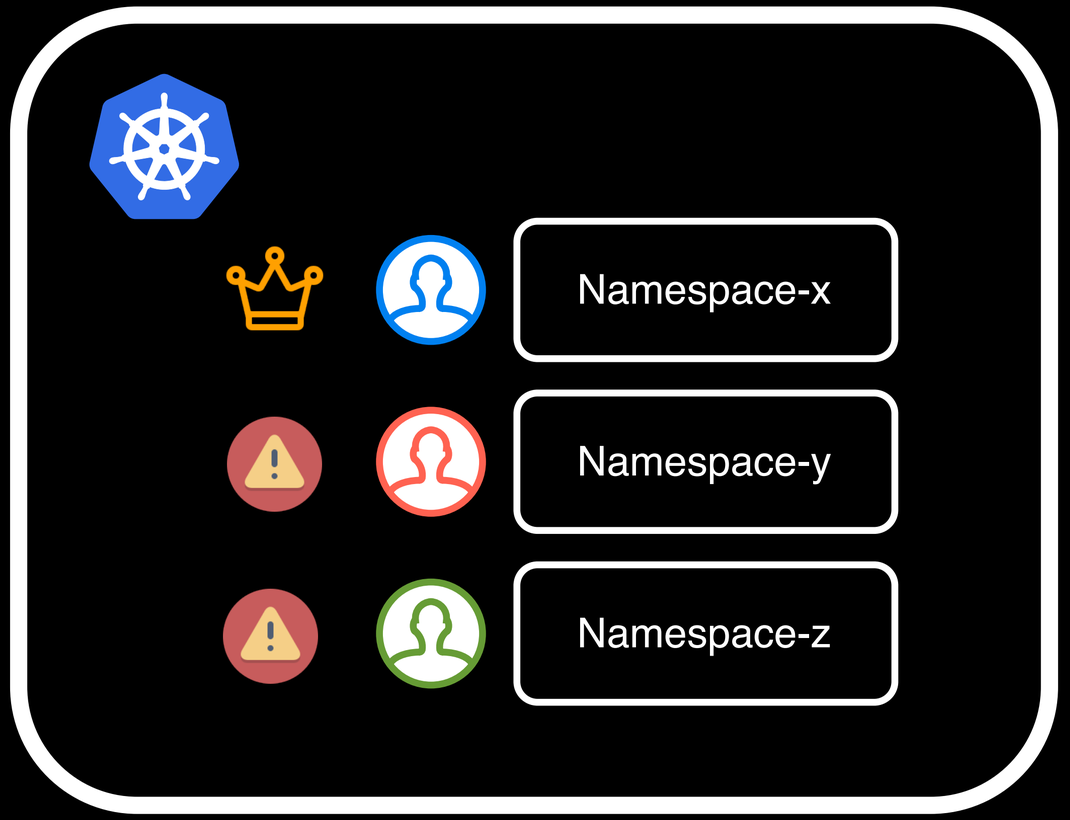
What you may see in real life



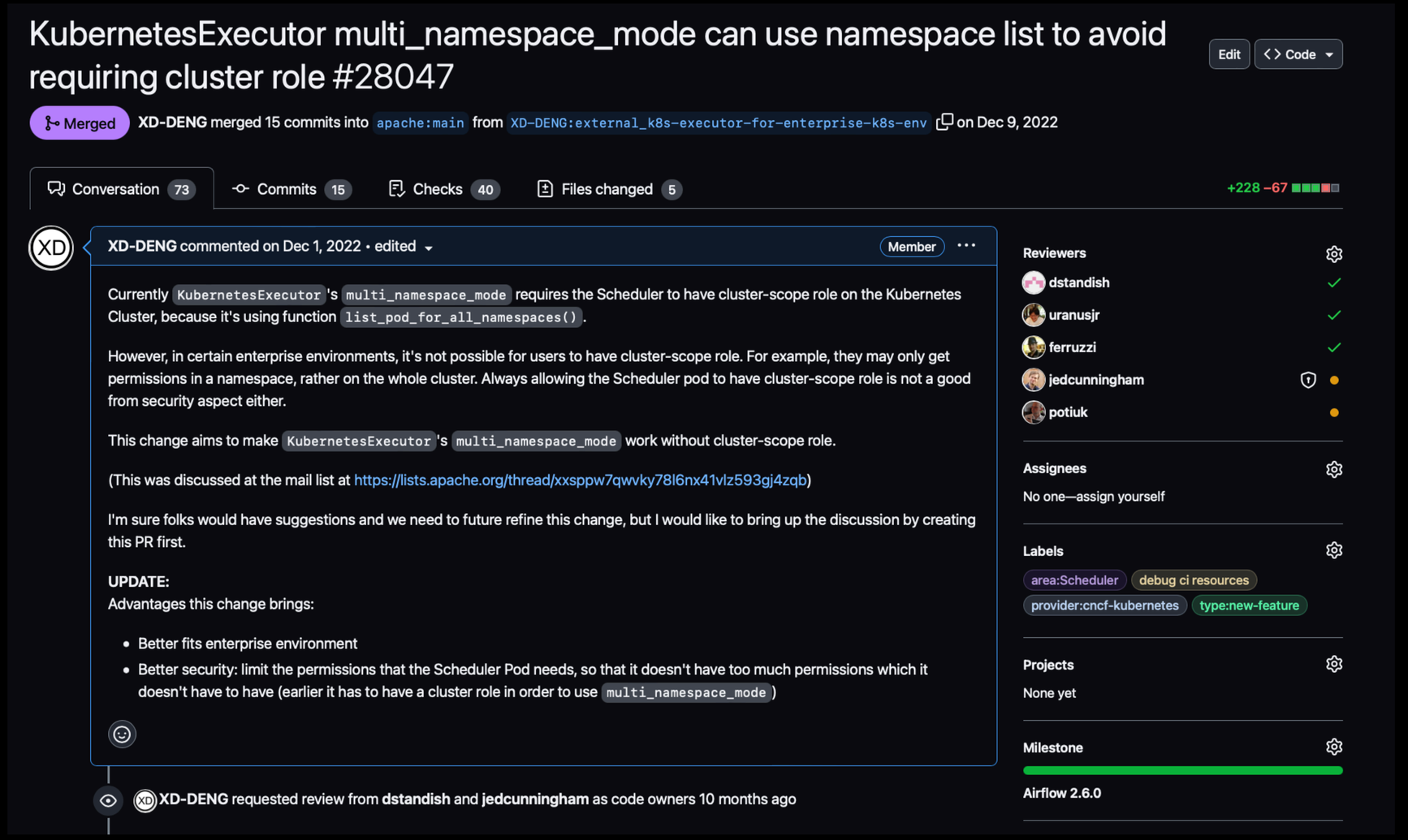
What was the pain/Why should there be a change?



What was the pain/Why should there be a change?



How does it work now?



KubernetesExecutor multi_namespace_mode can use namespace list to avoid requiring cluster role #28047

Edit · Code

Merged · XD-DENG merged 15 commits into `apache:main` from `XD-DENG:external_k8s-executor-for-enterprise-k8s-env` on Dec 9, 2022

Conversation 73 · Commits 15 · Checks 40 · Files changed 5 · +228 −67

XD-DENG commented on Dec 1, 2022 • edited · Member

Currently `KubernetesExecutor`'s `multi_namespace_mode` requires the Scheduler to have cluster-scope role on the Kubernetes Cluster, because it's using function `list_pod_for_all_namespaces()`.

However, in certain enterprise environments, it's not possible for users to have cluster-scope role. For example, they may only get permissions in a namespace, rather on the whole cluster. Always allowing the Scheduler pod to have cluster-scope role is not a good from security aspect either.

This change aims to make `KubernetesExecutor`'s `multi_namespace_mode` work without cluster-scope role.

(This was discussed at the mail list at https://lists.apache.org/thread/xxsppw7qwvky78l6nx41vlz593gj4zqb)

I'm sure folks would have suggestions and we need to future refine this change, but I would like to bring up the discussion by creating this PR first.

UPDATE:
Advantages this change brings:

- Better fits enterprise environment
- Better security: limit the permissions that the Scheduler Pod needs, so that it doesn't have too much permissions which it doesn't have to have (earlier it has to have a cluster role in order to use `multi_namespace_mode`)

XD-DENG requested review from **dstandish** and **jedcunningham** as code owners 10 months ago

Reviewers
dstandish
uranusjr
ferruzzi
jedcunningham
potiuk

Assignees
No one—assign yourself

Labels
area:Scheduler · debug ci resources · provider:cncf-kubernetes · type:new-feature

Projects
None yet

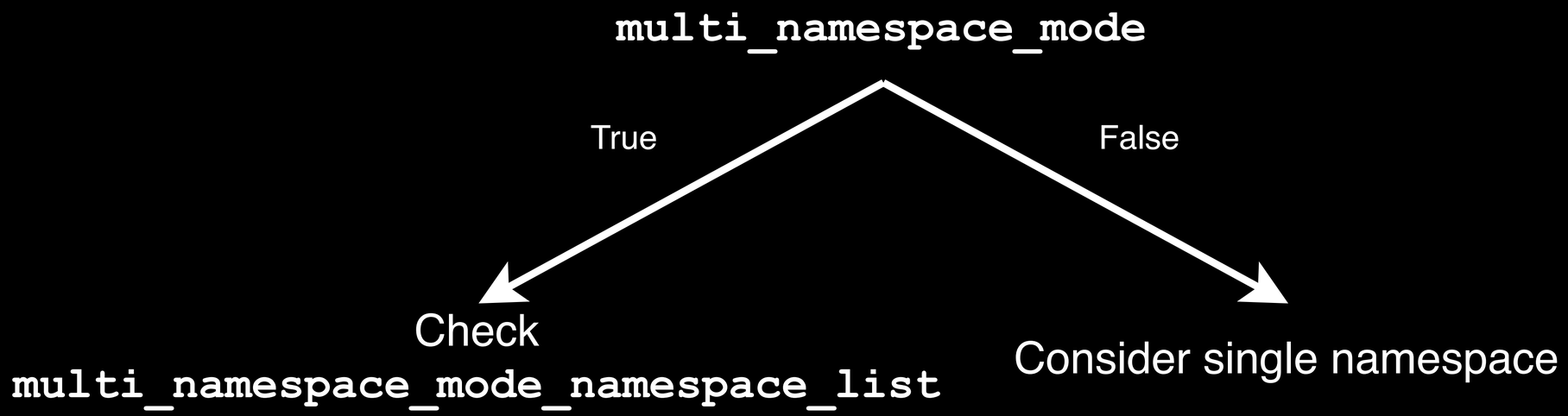
Milestone
Airflow 2.6.0

How does it work now?

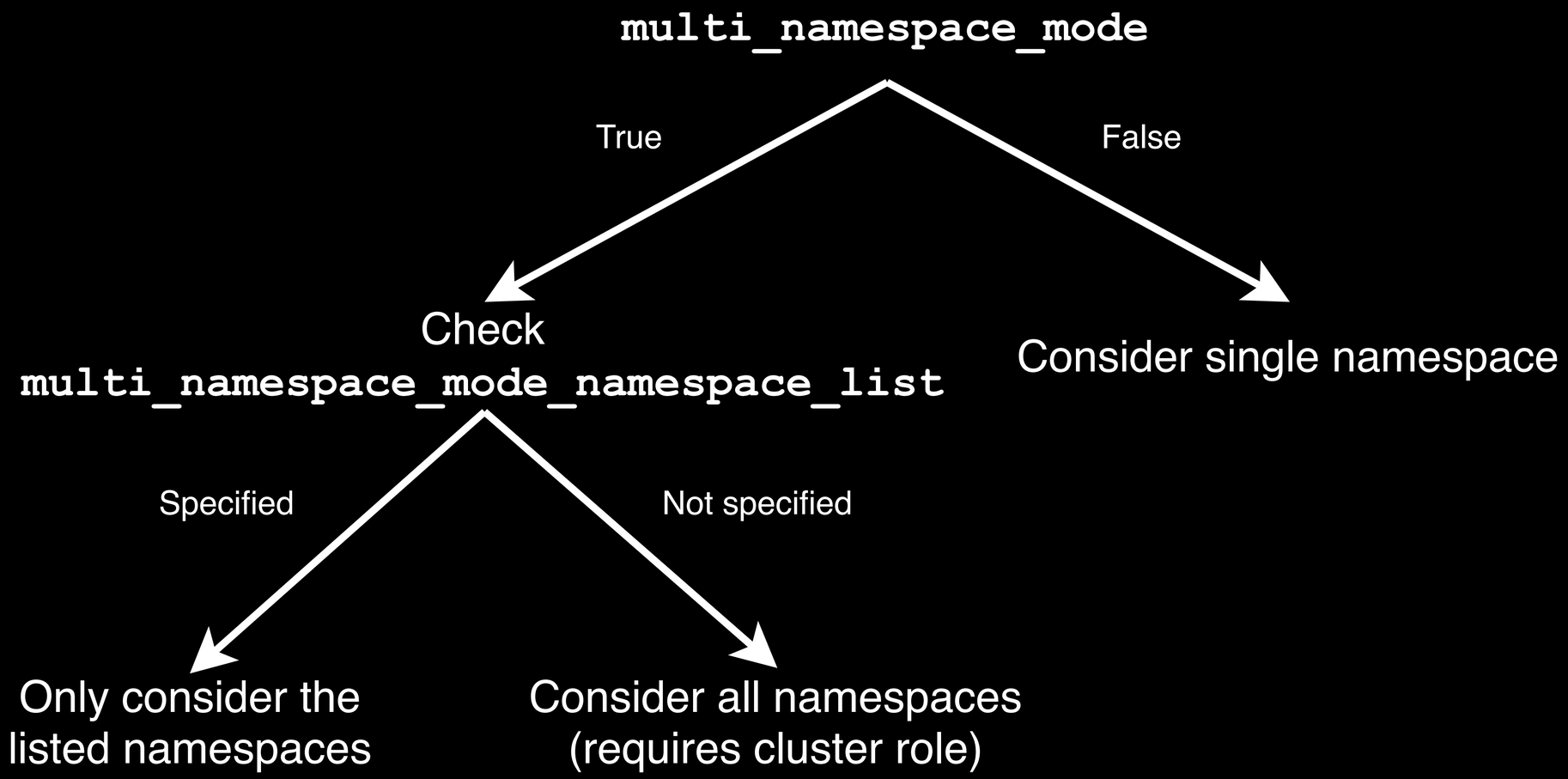
Airflow 2.6.0 & higher

```
# Allows users to launch pods in multiple namespaces.
# Will require creating a cluster-role for the scheduler,
# or use multi_namespace_mode_namespace_list configuration.
multi_namespace_mode = False

# If multi_namespace_mode is True while scheduler does not have a cluster-role,
# give the list of namespaces where the scheduler will schedule jobs
# Scheduler needs to have the necessary permissions in these namespaces.
multi_namespace_mode_namespace_list =
```

How does it work now?



How does it work now?



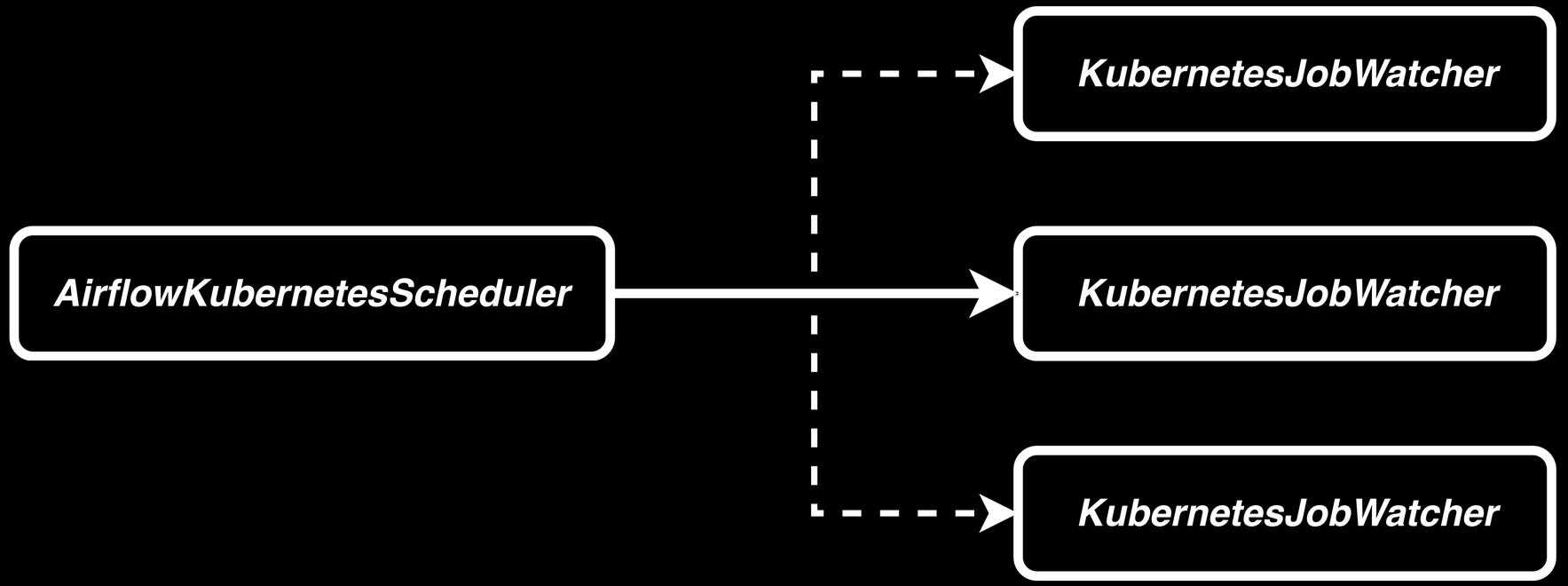
How does it work now?

An example

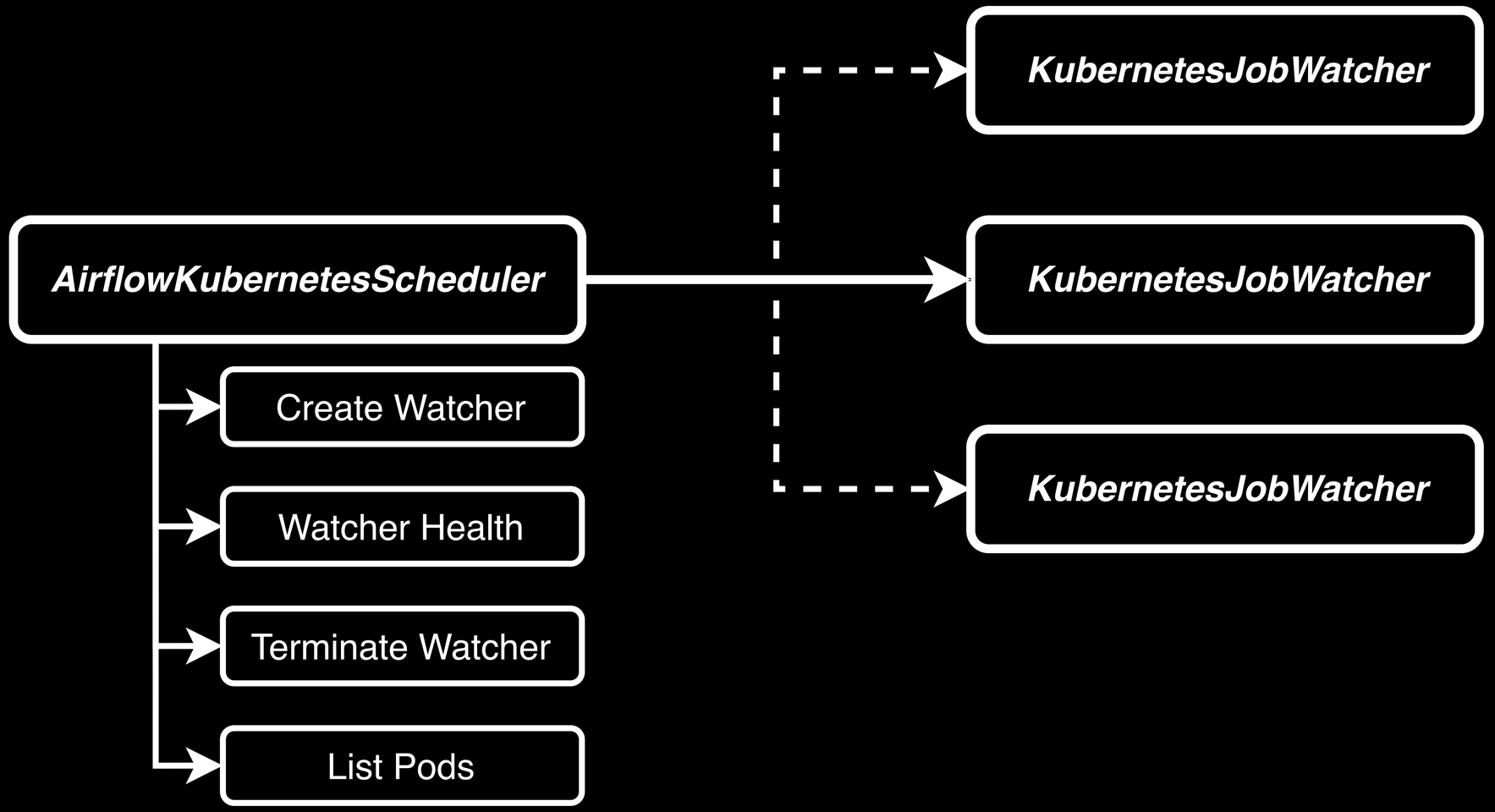
```
# Allows users to launch pods in multiple namespaces.
# Will require creating a cluster-role for the scheduler,
# or use multi_namespace_mode_namespace_list configuration.
multi_namespace_mode = True

# If multi_namespace_mode is True while scheduler does not have a cluster-role,
# give the list of namespaces where the scheduler will schedule jobs
# Scheduler needs to have the necessary permissions in these namespaces.
multi_namespace_mode_namespace_list = namespace_a,namespace_b,namespace_c
```

How does it work now?

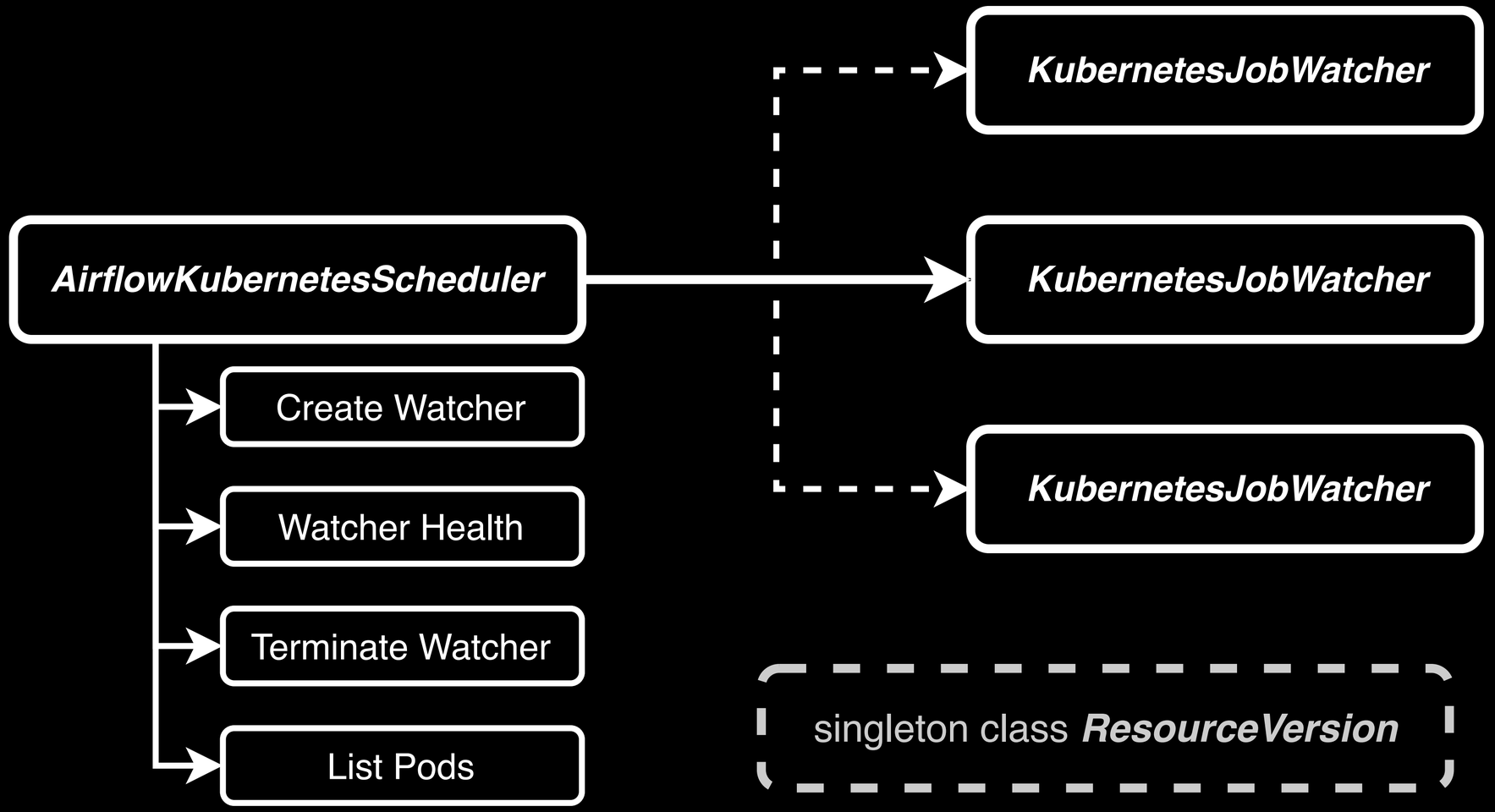
How does it work now?



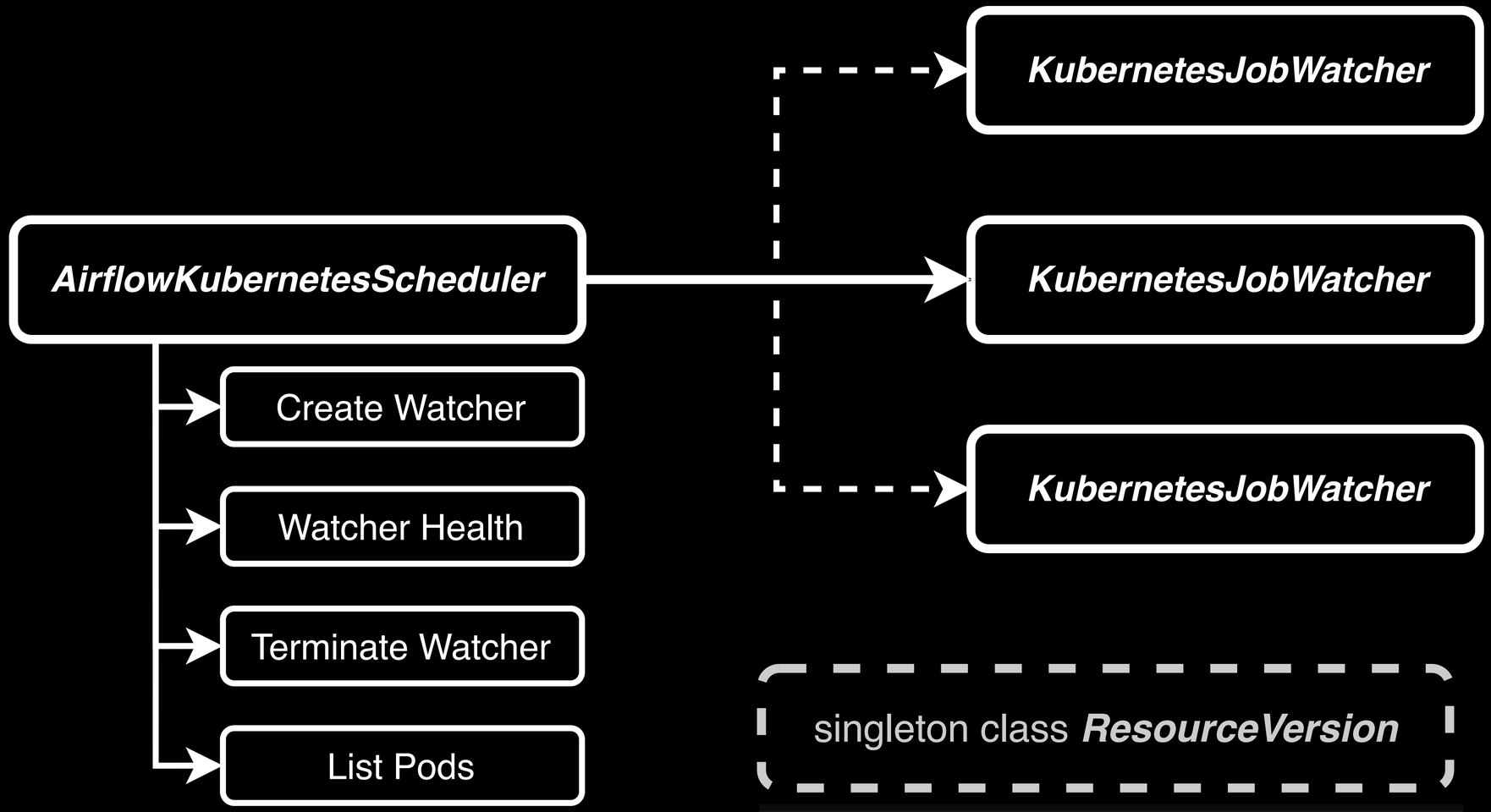
How does it work now?



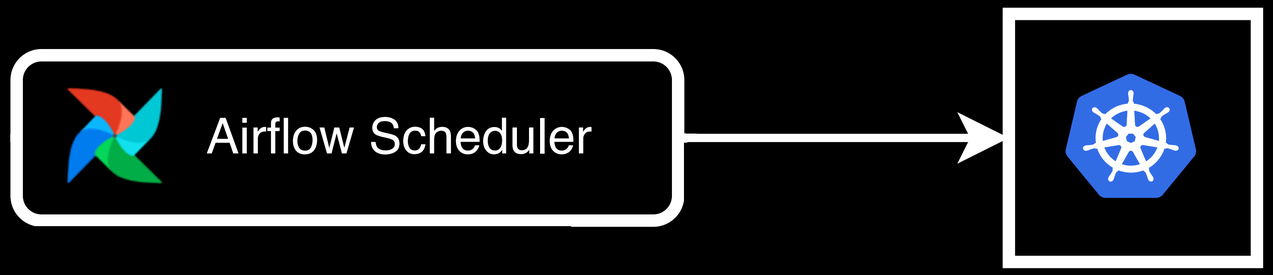
How does it work now?



How does it work now?

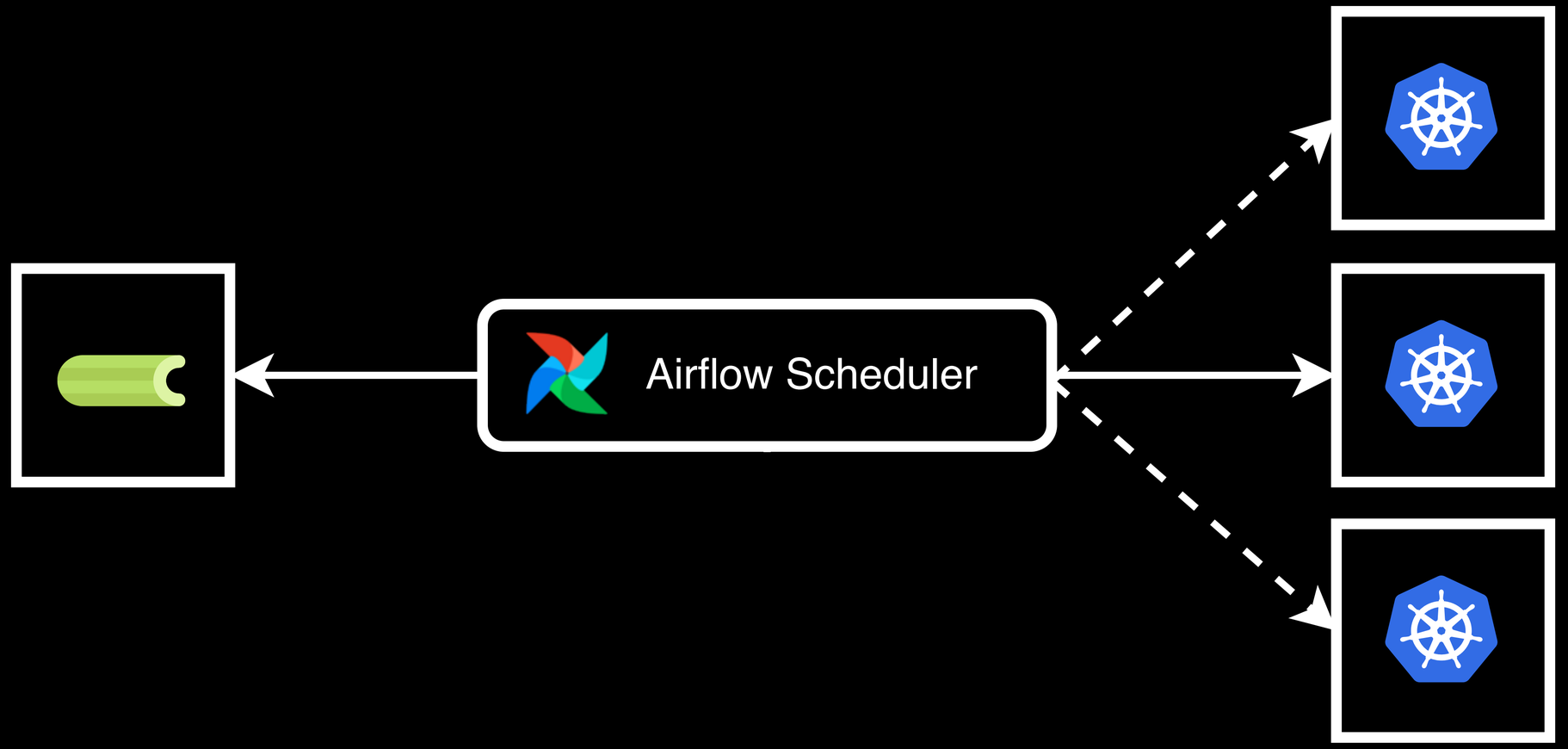


What can be the next improvement?

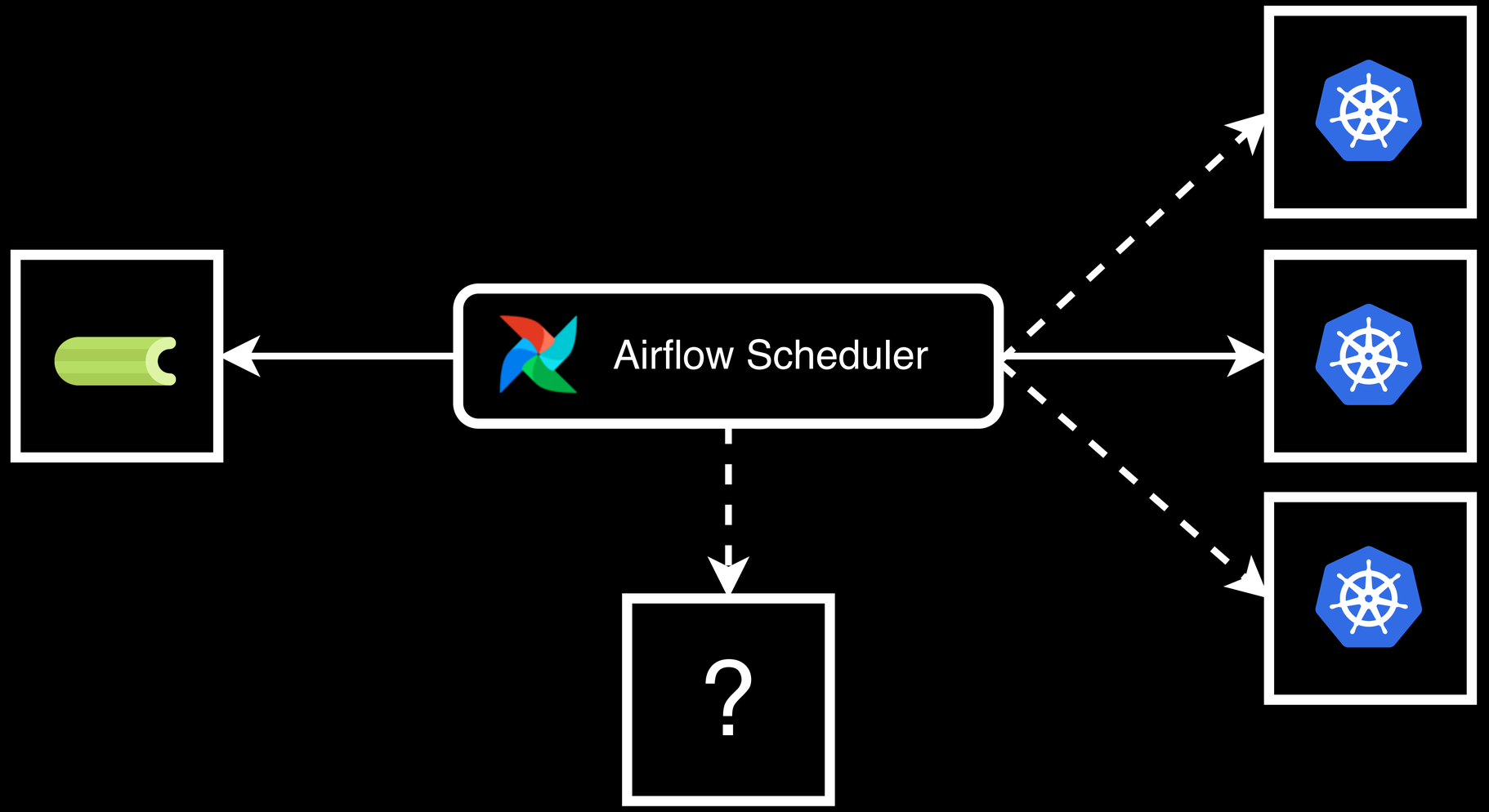


What can be the next improvement?

What can be the next improvement?



What can be the next improvement?



Thanks!