Hands-on Lab: Getting started with Apache Airflow

Estimated time needed: **20** minutes

Objectives

After completing this lab you will be able to:

- Start Apache Airflow.
- Open the Airflow UI in a browser.
- List all the DAGs.
- List the tasks in a DAG.
- Explore a DAG in the UI.

About Skills Network Cloud IDE

Skills Network Cloud IDE (based on Theia and Docker) provides an environment for hands on labs for course and project related labs. Theia is an open source IDE (Integrated Development Environment), that can be run on desktop or on the cloud. To complete this lab, we will be using the Cloud IDE based on Theia running in a Docker container.

Important Notice about this lab environment

Please be aware that sessions for this lab environment are not persistent. A new environment is created for you every time you connect to this lab. Any data you may have saved in an earlier session will get lost. To avoid losing your data, please plan to complete these labs in a single session.

Exercise 1 - Start Apache Airflow

Open a new terminal by clicking on the menu bar and selecting **Terminal->New Terminal**, as shown in the image below.
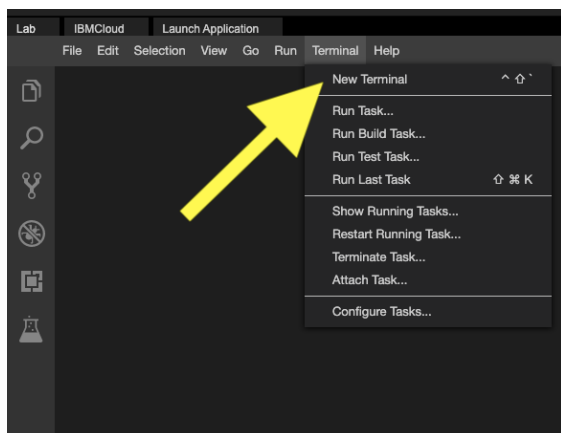
This will open a new terminal at the bottom of the screen as in the image below.
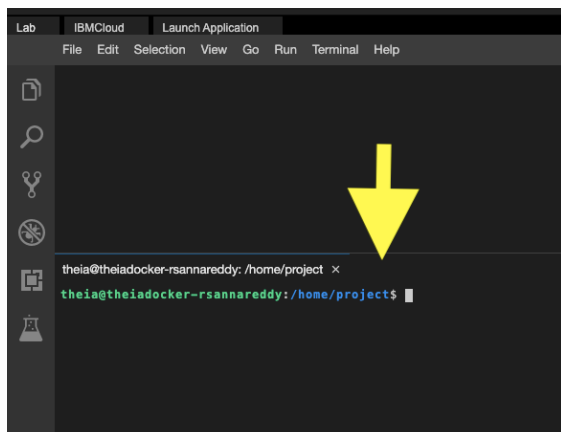
Run the commands below on the newly opened terminal. (You can copy the code by clicking on the little copy button on the bottom right of the codeblock below and paste it wherever you wish.)

Run the command below in the terminal to start Apache Airflow.

1. 1

```
1. start_airflow
```

Copied!

Please be patient, it may take upto 5 minutes for airflow to get started.

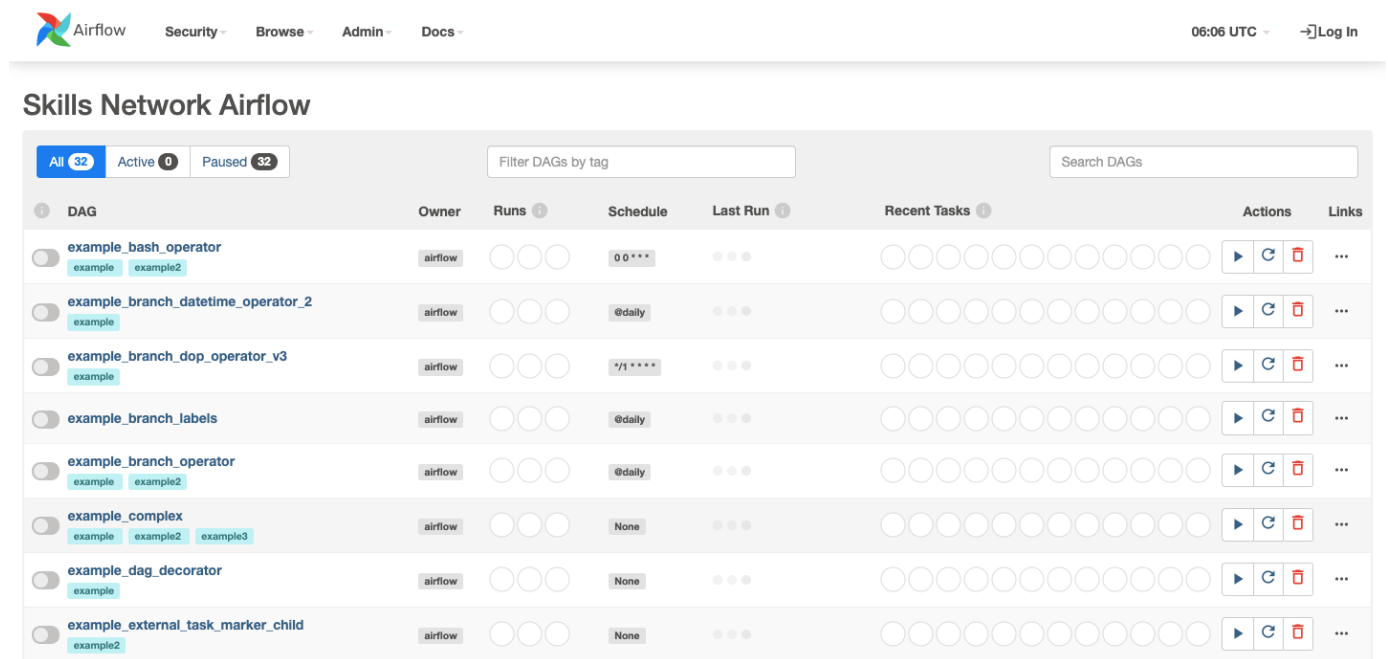When airflow starts successfully, you should see an output similar to the one below.
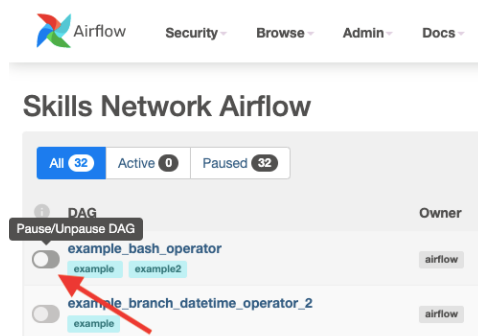
Exercise 2 - Open the Airflow Web UI

Copy the Web-UI URL and paste it on a new browser tab. You can also click on the URL by holding the control key (Command key in case of a Mac).

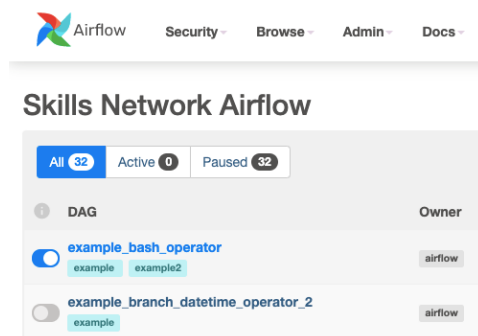You should land at a page that looks like this:

You can Unpause/Pause a DAG using the Unpause/Pause toggle button.

An unpaused DAG looks like this:

Click on a DAG to explore more about the DAG.
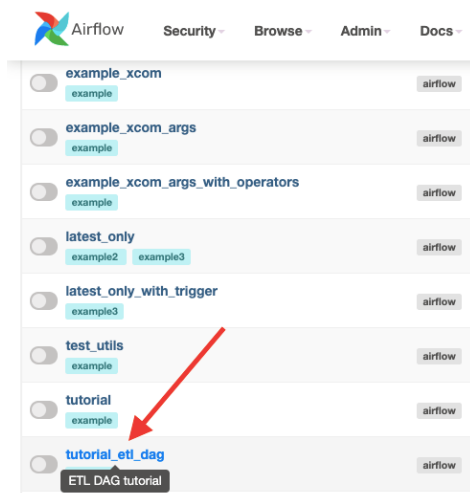
Click on the Grid button to see the grid view of the DAG.
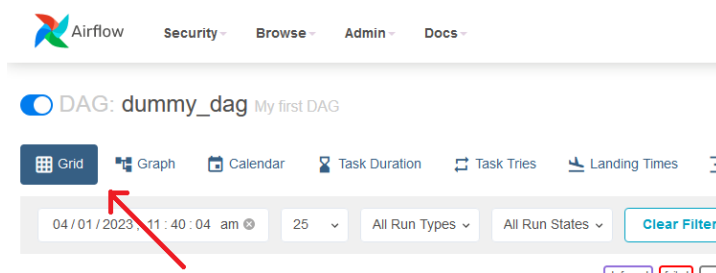
Click on the Graph button to see the graph view of the DAG.

Exercise 3 - List all DAGs

Apache airflow gives us some handy command line options to work with.

Run the command below in the terminal to list out all the existing DAGs.

```
1. 1
1. airflow dags list
```

Copied!

Exercise 4 - List tasks in a DAG

Run the command below in the terminal to list out all the tasks in the DAG named `example_bash_operator`.

```
1. 1
1. airflow tasks list example_bash_operator
```

Copied!

Here `example_bash_operator` is the name of the DAG.

Try listing out the tasks for the DAG `tutorial`.

Exercise 5 - Pause/Unpause a DAG

Run the command below in the terminal to unpause a DAG.

```
1. 1
1. airflow dags unpause tutorial
```

Copied!

Here `tutorial` is the name of the DAG.

Pause a DAG.

```
1. 1
1. airflow dags pause tutorial
```

Copied!

Try to unpause the DAG named `example_branch_operator`.

Practice exercises

1. Problem:

   List tasks for the DAG `example_branch_labels`.

▸ Click here for Hint

▾ Click here for Solution

1. 1

   1. `airflow tasks list example_branch_labels`

Copied!

2. Problem:

   Unpause the DAG example_branch_labels.

▸ Click here for Hint
▾ Click here for Solution

1. 1

   1. `airflow dags unpause example_branch_labels`

Copied!

3. Problem:

   Pause the DAG example_branch_labels.

▸ Click here for Hint
▾ Click here for Solution

1. 1

   1. `airflow dags pause example_branch_labels`

Copied!

Authors

Ramesh Sannareddy

Other Contributors

Rav Ahuja

Change Log

Date (YYYY-MM-DD)	Version	Changed By	Change Description
2023-05-04	0.2	Benny Li	Minor formatting changes
2021-07-05	0.1	Ramesh Sannareddy	Created initial version of the lab