

# Hands-on Lab: Getting Started with Apache Airflow



Estimated time needed: **20** minutes

## Introduction

In this lab, you will have the opportunity to work with Apache Airflow. You will work with both the Web user interface (UI) and command line interface (CLI) to explore the various ways you can use Airflow for workflow management.

## 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

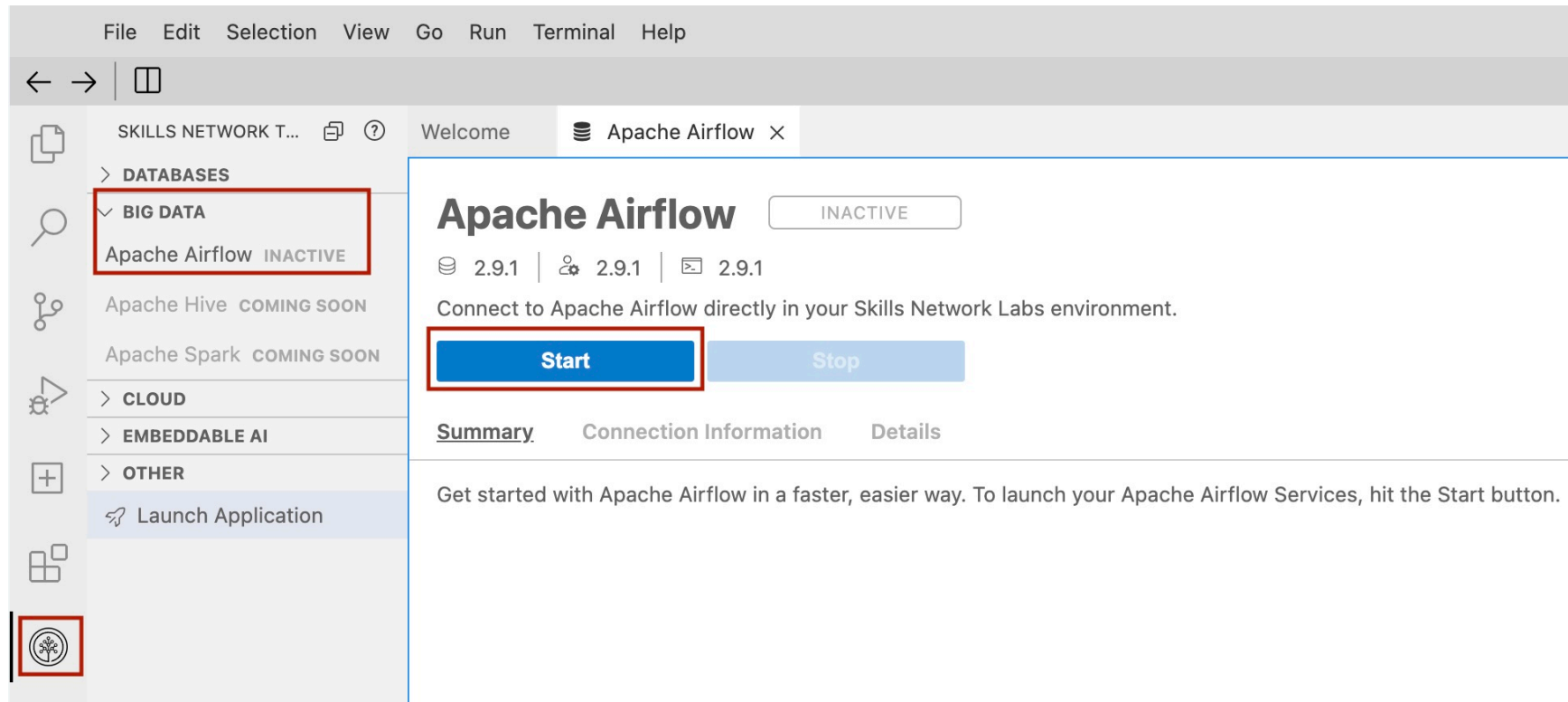
Skills Network Cloud IDE (based on Theia and Docker) provides an environment for hands on labs for course and project related labs. Theia is an open source IDE (Integrated Development Environment), that can be run on desktop or on the cloud. To complete this lab, we will be using the Cloud IDE based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

1. Click on **Skills Network Toolbox**.
2. From the **BIG DATA** section, click **Apache Airflow**.
3. Click **Start** to start the Apache Airflow.



**Note:** Please be patient, it will take a few minutes for Airflow to start.

# Exercise 2: Open the Airflow Web UI

1. When Airflow starts successfully, you should see an output similar to the one below. Once **Apache Airflow** has started, click on the highlighted icon to open **Apache Airflow Web UI** in the new window.

# Apache Airflow

**ACTIVE** 2.9.1 |  2.9.1 |  2.9.1

Connect to Apache Airflow directly in your Skills Network Labs environment.

**Start****Stop****Summary**

Connection Information

Details

Your Apache Airflow Services are now ready to use and available with the following login credentials. For more details on how to navigate Apache Airflow, please check out the Details section.

**Username:**

airflow

**Password:**

MzE3NjUtbGF2YW55

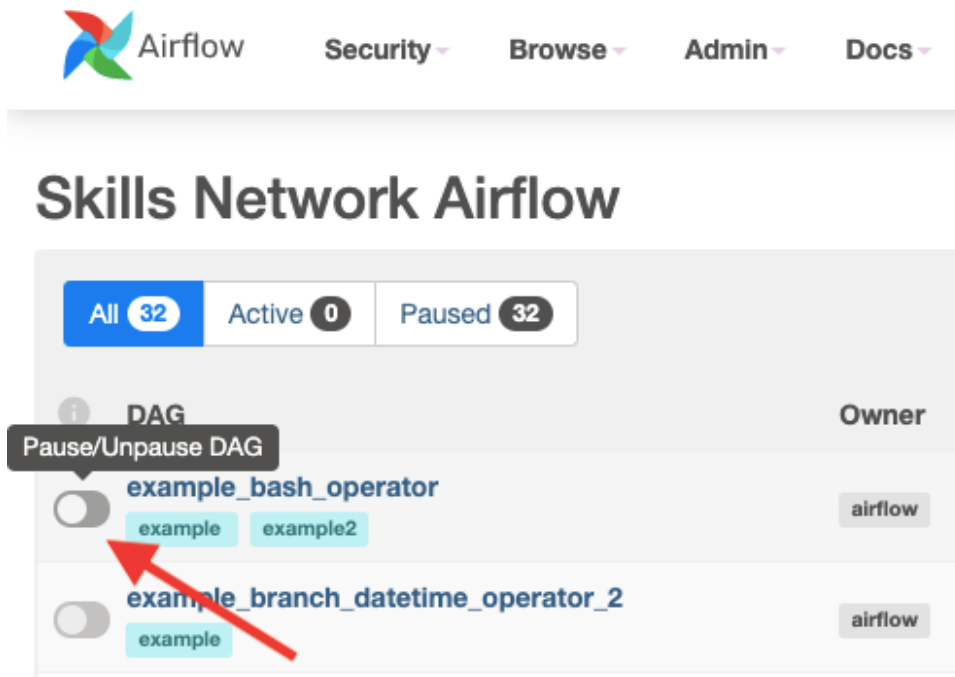


You can manage Apache Airflow via:

**Airflow Webserver**

2. You will land on a page that displays various DAGs and associated options.

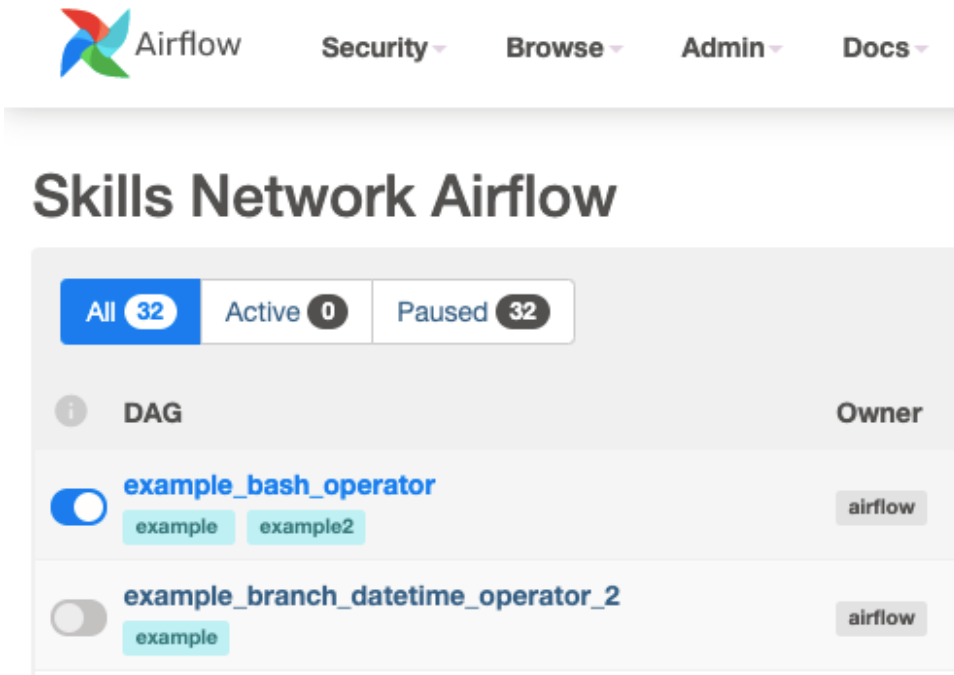




The image shows the Airflow Skills Network interface. At the top, there is a navigation bar with the Airflow logo and links for Security, Browse, Admin, and Docs. Below this, the title "Skills Network Airflow" is displayed. A filter bar shows "All 32", "Active 0", and "Paused 32". The main content area is a table with columns for DAG, Owner, and a toggle button. The first row shows "example\_bash\_operator" with owner "airflow" and a toggle button. A tooltip "Pause/Unpause DAG" is visible over the toggle button. The second row shows "example\_branch\_datetime\_operator\_2" with owner "airflow" and a toggle button. A red arrow points to the toggle button of the second row.

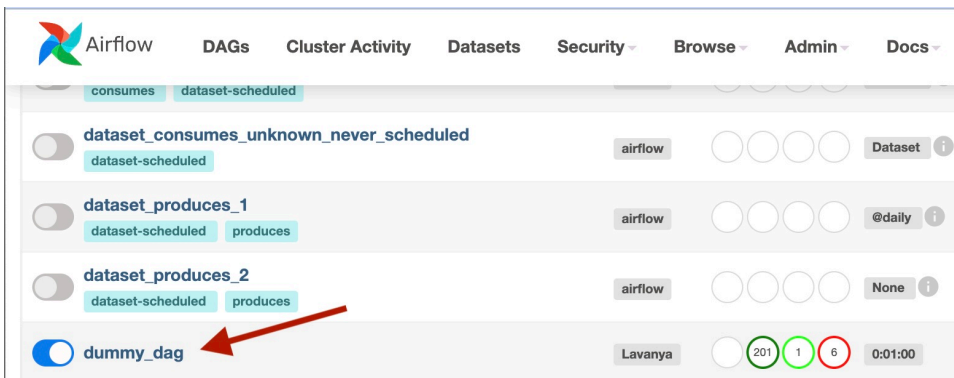
DAG	Owner
<input type="checkbox"/> example_bash_operator example example2	airflow
<input type="checkbox"/> example_branch_datetime_operator_2 example	airflow

The button is not greyed out when the DAG is running.



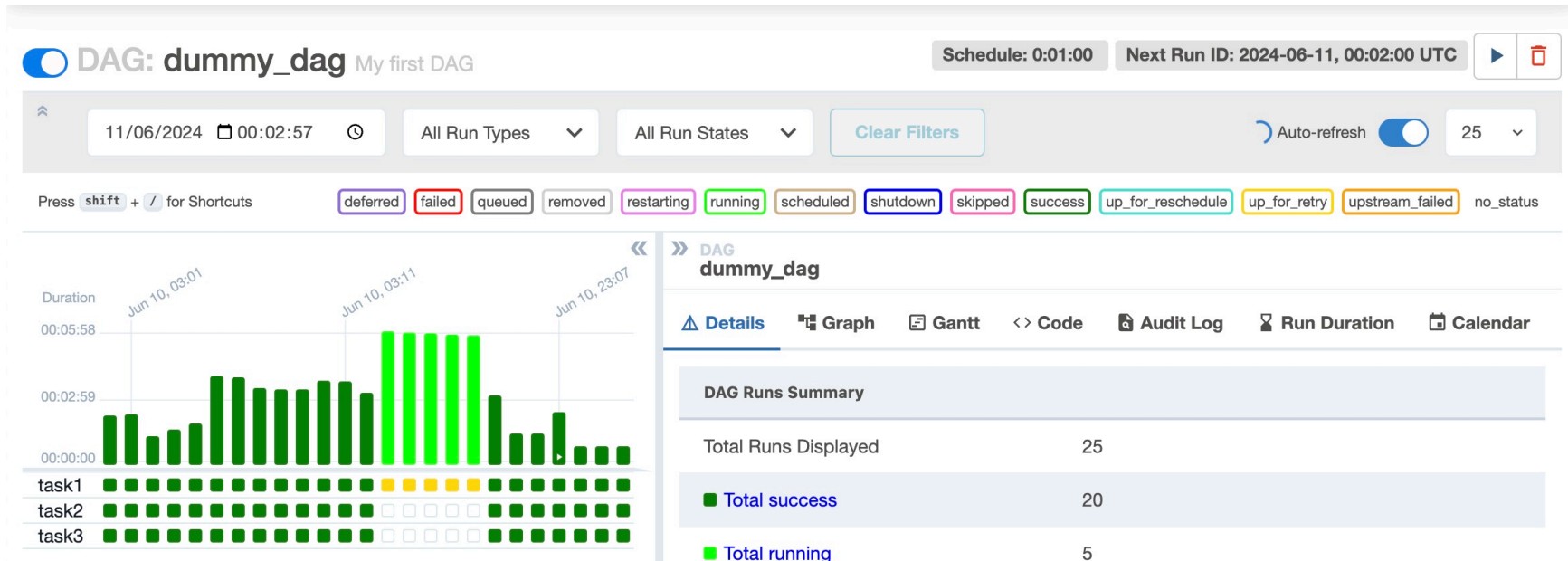
The image shows the Airflow Skills Network interface. At the top, there is a navigation bar with the Airflow logo and links for Security, Browse, Admin, and Docs. Below the navigation bar, the title "Skills Network Airflow" is displayed. Underneath the title, there are three tabs: "All 32", "Active 0", and "Paused 32". The "All 32" tab is selected. Below the tabs, there is a table with two columns: "DAG" and "Owner". The table lists two DAGs: "example\_bash\_operator" and "example\_branch\_datetime\_operator\_2". The "example\_bash\_operator" DAG has a toggle switch that is turned on and shows two tags: "example" and "example2". The "example\_branch\_datetime\_operator\_2" DAG has a toggle switch that is turned off and shows one tag: "example". Both DAGs have "airflow" as the owner.

4. Click on a DAG to explore more.

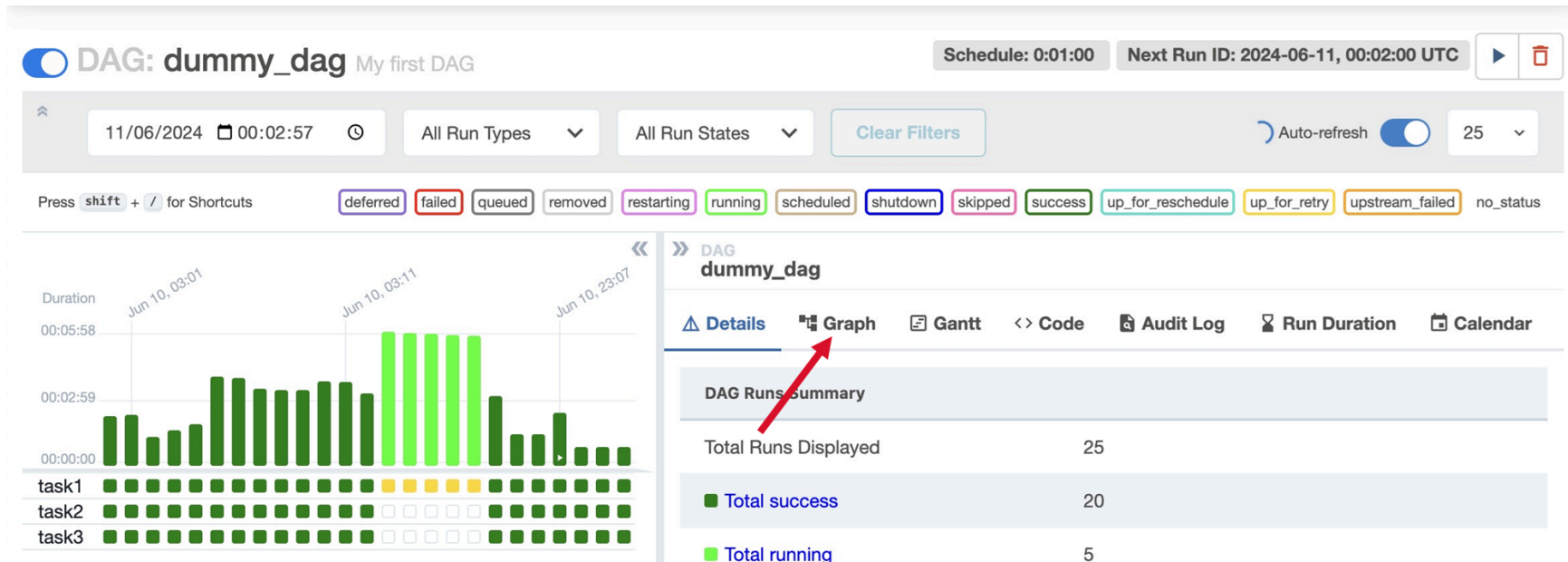


The image shows the Airflow DAGs list interface. At the top, there is a navigation bar with the Airflow logo and links for DAGs, Cluster Activity, Datasets, Security, Browse, Admin, and Docs. Below the navigation bar, there is a table with four columns: "DAG", "Owner", "Status", and "Frequency". The table lists four DAGs: "dataset\_consumes\_unknown\_never\_scheduled", "dataset\_produces\_1", "dataset\_produces\_2", and "dummy\_dag". The "dataset\_consumes\_unknown\_never\_scheduled" DAG has a toggle switch that is turned off and shows one tag: "dataset-scheduled". The "dataset\_produces\_1" DAG has a toggle switch that is turned off and shows two tags: "dataset-scheduled" and "produces". The "dataset\_produces\_2" DAG has a toggle switch that is turned off and shows two tags: "dataset-scheduled" and "produces". The "dummy\_dag" DAG has a toggle switch that is turned on and shows no tags. A red arrow points to the "dummy\_dag" toggle switch. The "dummy\_dag" DAG has "Lavanya" as the owner, a status of "201 1 6" (with 201 in a green circle, 1 in a green circle, and 6 in a red circle), and a frequency of "0:01:00".

5. You will see the Grid view by default.

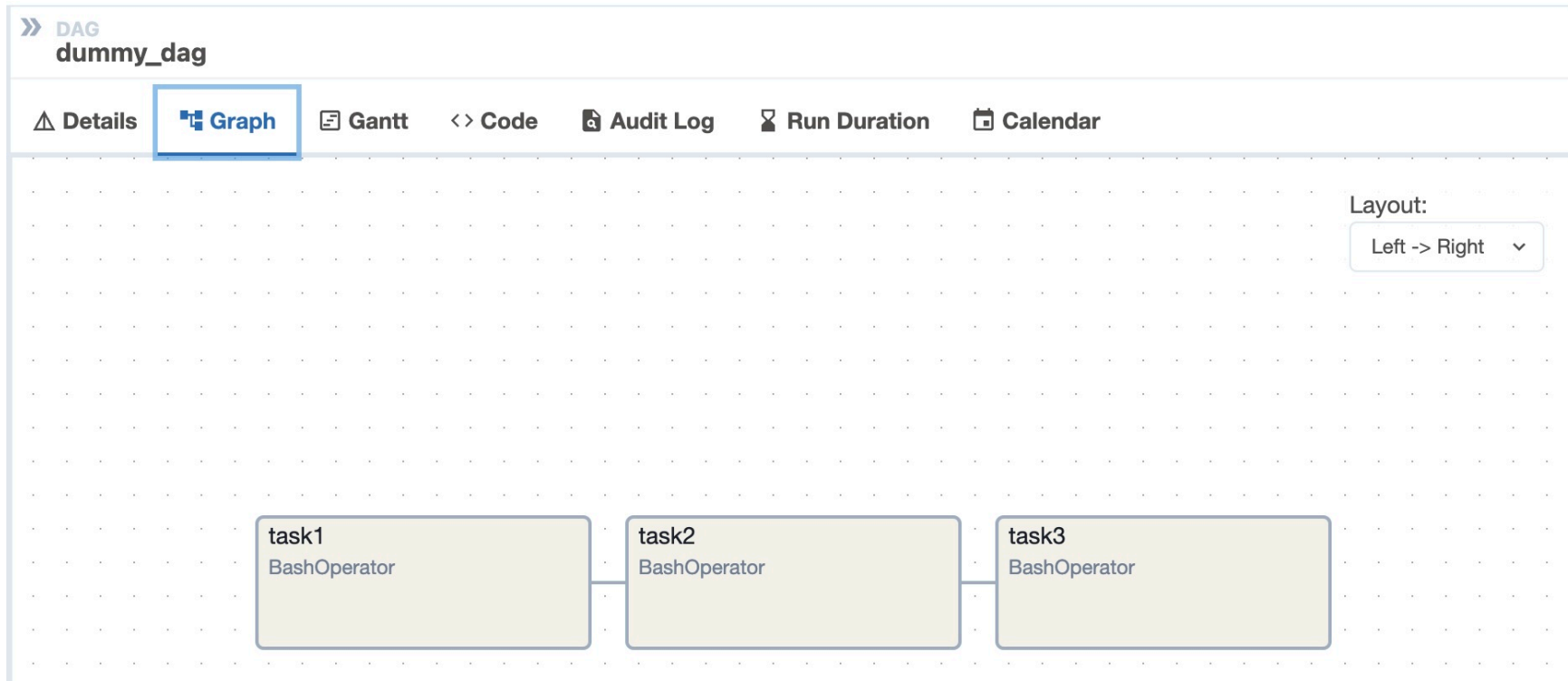


6. Next, click on the **Graph** button.



7. Notice the graph view of the DAG will display.





## Exercise 3: Apache Airflow CLI

Apache Airflow provides some command line options.

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

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

Copied! Executed!

2. Run the command below in the terminal to list all tasks in the DAG named example\_bash\_operator.

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

Copied! Executed!

## Practice Exercise

1. Run a command to list all tasks for the DAG named tutorial.

▼ [Click here for the solution.](#)

```
1. 1
```

```
1. airflow tasks list tutorial
```

Copied!

Executed!

## Exercise 4: Pause or Unpause a DAG

1. Run the command below in the terminal to unpause a DAG named tutorial.

```
1. 1
```

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

Copied!

2. Run the command to pause the DAG.

```
1. 1
```

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

Copied!

## Practice Exercise

1. Run a command to unpause the DAG named example\_branch\_operator.

▼ [Click here for the solution.](#)

```
1. 1
```

```
1. airflow dags unpause example_branch_operator
```

Copied!

## Practice exercises

1. List tasks for the DAG example\_branch\_labels.

► [Click here for hint.](#)

▼ [Click here for the solution.](#)

1. 1

1. airflow tasks list example\_branch\_labels

Copied!

2. Unpause the DAG example\_branch\_labels.

► [Click here for hint.](#)

▼ [Click here for the solution.](#)

1. 1

1. airflow dags unpause example\_branch\_labels

Copied!

3. Pause the DAG example\_branch\_labels.

► [Click here for hint.](#)

▼ [Click here for the solution.](#)

1. 1

1. airflow dags pause example\_branch\_labels

Copied!

## Authors

Ramesh Sannareddy

[Lavanya T S](#)

## Other Contributors

Rav Ahuja

© IBM Corporation. All rights reserved.