

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

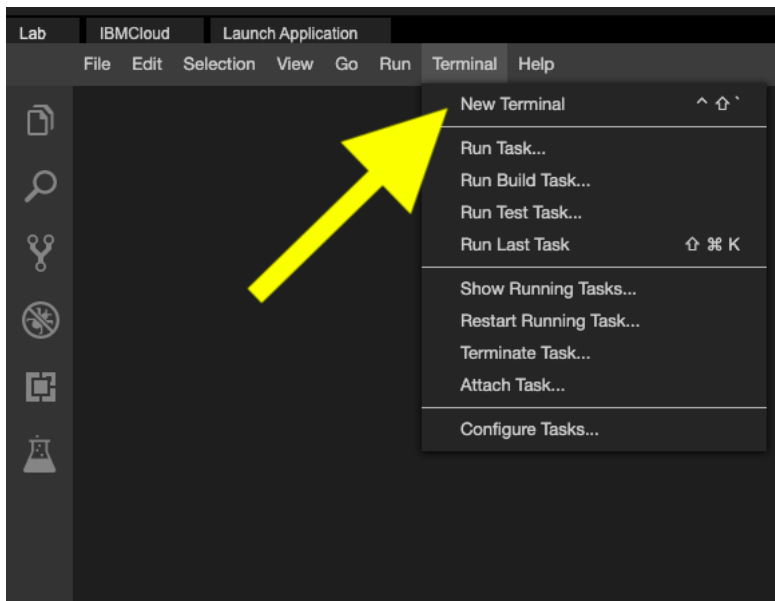
Skills Network Cloud IDE (based on Theia and Docker) provides an environment for hands on labs for course and project related labs. Theia is an open source IDE (Integrated Development Environment), that can be run on desktop or on the cloud. To complete this lab, we will be using the Cloud IDE based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

Important Notice about this lab environment

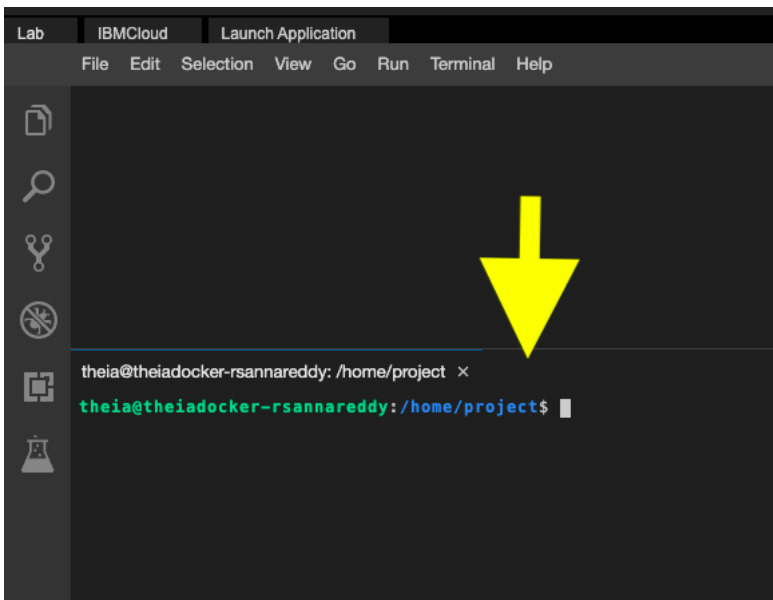
Please be aware that sessions for this lab environment are not persistent. A new environment is created for you every time you connect to this lab. Any data you may have saved in an earlier session will get lost. To avoid losing your data, please plan to complete these labs in a single session.

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.

```
theia@theiadocker-rsannareddy:/home/project$ start_airflow
Starting your airflow services....
This process can take a few minutes.

Airflow started, waiting for all services to be ready....

Your airflow server is now ready to use and available with username: airflow password: MTM4ODUtcnNhbm5h

You can access your Airflow Webserver at: https://rsannareddy-8080.theiadocker-5-labs-prod-th
eiak8s-4-tor01.proxy.cognitiveclass.ai

CommandLine:
• List DAGs: airflow dags list
• List Tasks: airflow tasks list example_bash_operator
• Run an example task: airflow tasks test example_bash_operator runme_1 2015-06-01
theia@theiadocker-rsannareddy:/home/project$
```

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:

Skills Network Airflow

<div> All 32 Active 0 Paused 32 </div> <div>Filter DAGs by tag</div>						
DAG	Owner	Runs	Schedule	Last Run	Recent Tasks	
<input type="checkbox"/> example_bash_operator example example2	airflow	<div><div></div><div></div><div></div></div>	00***	<div><div></div><div></div><div></div></div>	<div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div>	
<input type="checkbox"/> example_branch_datetime_operator_2 example	airflow	<div><div></div><div></div><div></div></div>	@daily	<div><div></div><div></div><div></div></div>	<div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div>	
<input type="checkbox"/> example_branch_dop_operator_v3 example	airflow	<div><div></div><div></div><div></div></div>	*/1****	<div><div></div><div></div><div></div></div>	<div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div>	
<input type="checkbox"/> example_branch_labels 	airflow	<div><div></div><div></div><div></div></div>	@daily	<div><div></div><div></div><div></div></div>	<div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div>	
<input type="checkbox"/> example_branch_operator example example2	airflow	<div><div></div><div></div><div></div></div>	@daily	<div><div></div><div></div><div></div></div>	<div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div>	
<input type="checkbox"/> example_complex example example2 example3	airflow	<div><div></div><div></div><div></div></div>	None	<div><div></div><div></div><div></div></div>	<div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div>	
<input type="checkbox"/> example_dag_decorator example	airflow	<div><div></div><div></div><div></div></div>	None	<div><div></div><div></div><div></div></div>	<div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div>	
<input type="checkbox"/> example_external_task_marker_child example2	airflow	<div><div></div><div></div><div></div></div>	None	<div><div></div><div></div><div></div></div>	<div><div></div><div></div><div></div><div></div><div></div><div></div><div></div><div></div></div>	

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

Skills Network Airflow

<div> All 32 Active 0 Paused 32 </div>						
DAG	Owner					
<input type="checkbox"/> example_bash_operator example example2	airflow					
<input type="checkbox"/> example_branch_datetime_operator_2 example	airflow					

An unpaused DAG looks like this:

Skills Network Airflow

<div> All 32 Active 0 Paused 32 </div>		
ⓘ DAG	Owner	
<input checked="" type="checkbox"/> example_bash_operator example example2	airflow	
<input type="checkbox"/> example_branch_datetime_operator_2 example	airflow	

Click on a DAG to explore more about the DAG.

<input type="checkbox"/> example_xcom example	airflow
<input type="checkbox"/> example_xcom_args example	airflow
<input type="checkbox"/> example_xcom_args_with_operators example	airflow
<input type="checkbox"/> latest_only example2 example3	airflow
<input type="checkbox"/> latest_only_with_trigger example3	airflow
<input type="checkbox"/> test_utils example	airflow
<input type="checkbox"/> tutorial example	airflow
<input type="checkbox"/> tutorial_etl_dag ETL DAG tutorial	airflow

Click on the Grid button to see the grid view of the DAG.

☒ DAG: dummy_dag My first DAG

☒ Grid
 ☐ Graph
 ☐ Calendar
 ☐ Task Duration
 ☐ Task Tries
 ☐ Landing Times

deferred
failed
running

Click on the Graph button to see the graph view of the DAG.

queued running success failed up_for_retry up_for_reschedule upstream_failed skipped scheduled no_status

Auto-refresh



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

Use `list` option.

▼ 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

Use the unpause option.

▼ 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

Use the pause option.

▼ 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

Copyright (c) 2023 IBM Corporation. All rights reserved.