

# Laboratório Prático: Monitorando um DAG



Tempo estimado necessário: **20 minutos**

## Introdução

Neste laboratório, você trabalhará com a interface web do Airflow e o CLI para explorar mais os DAGs. Você será exposto ao uso de ferramentas interativas para buscar DAGs, conhecer várias visualizações dos DAGs e como você pode usar isso para explorar o fluxo de trabalho do DAG, as tarefas individuais no fluxo de trabalho e ver o resultado das tarefas.

## Objetivos

Após concluir este laboratório, você será capaz de:

- Buscar um DAG
- Pausar/Despausar um DAG
- Obter os Detalhes de um DAG
- Explorar a visualização em grade de um DAG
- Explorar a visualização em gráfico de um DAG
- Explorar a visualização em calendário de um DAG
- Explorar a visualização de Duração da Tarefa de um DAG
- Explorar a visualização de Detalhes de um DAG
- Ver o código-fonte de um DAG
- Deletar um DAG

## Sobre o Skills Network Cloud IDE

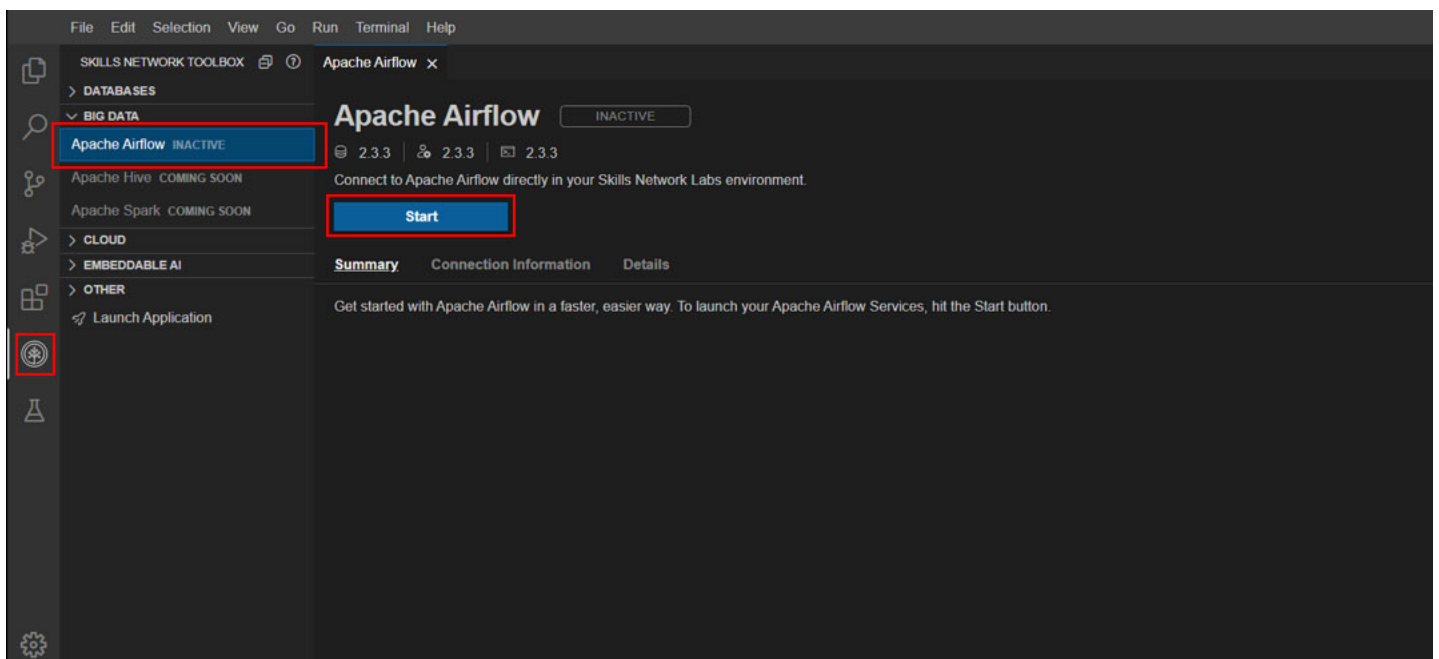
O Skills Network Cloud IDE (baseado no Theia e Docker) fornece um ambiente para laboratórios práticos relacionados a cursos e projetos. O Theia é um IDE (Ambiente de Desenvolvimento Integrado) de código aberto, que pode ser executado no desktop ou na nuvem. Para completar este laboratório, usaremos o Cloud IDE baseado no Theia rodando em um contêiner Docker.

## Aviso importante sobre este ambiente de laboratório

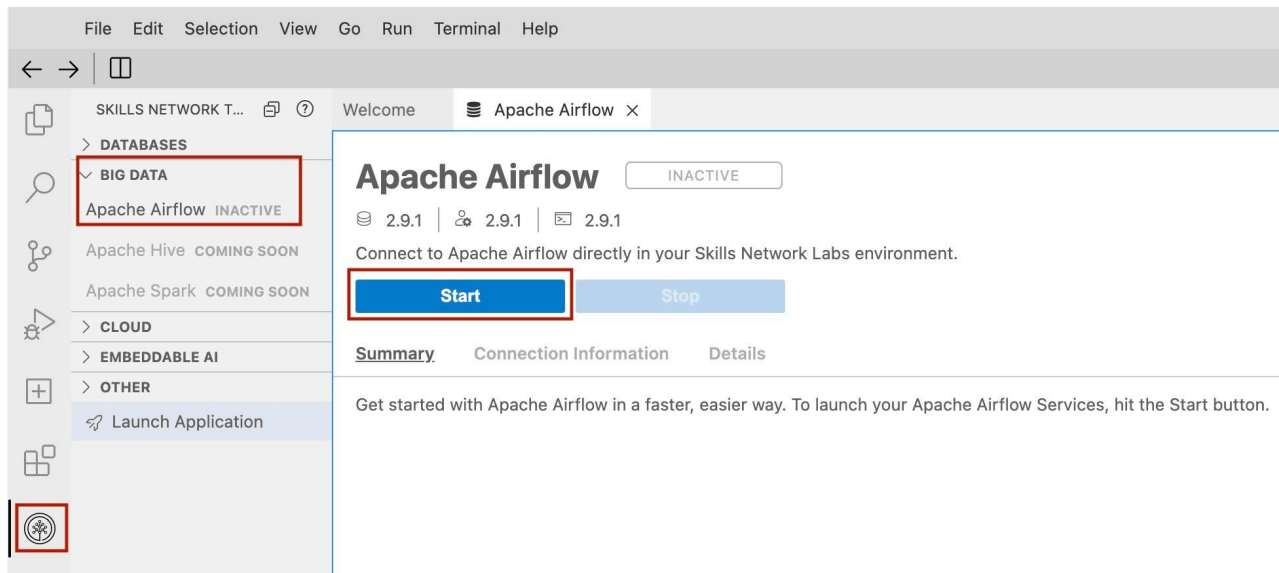
Por favor, esteja ciente de que as sessões para este ambiente de laboratório não são persistentes. Um novo ambiente é criado para você toda vez que você se conecta a este laboratório. Qualquer dado que você possa ter salvo em uma sessão anterior será perdido. Para evitar a perda de seus dados, planeje completar estes laboratórios em uma única sessão.

## Exercício 1: Iniciar o Apache Airflow

1. Clique em **Skills Network Toolbox**.
2. Na seção **BIG DATA**, clique em **Apache Airflow**.
3. Clique em **Iniciar** para iniciar o Apache Airflow.

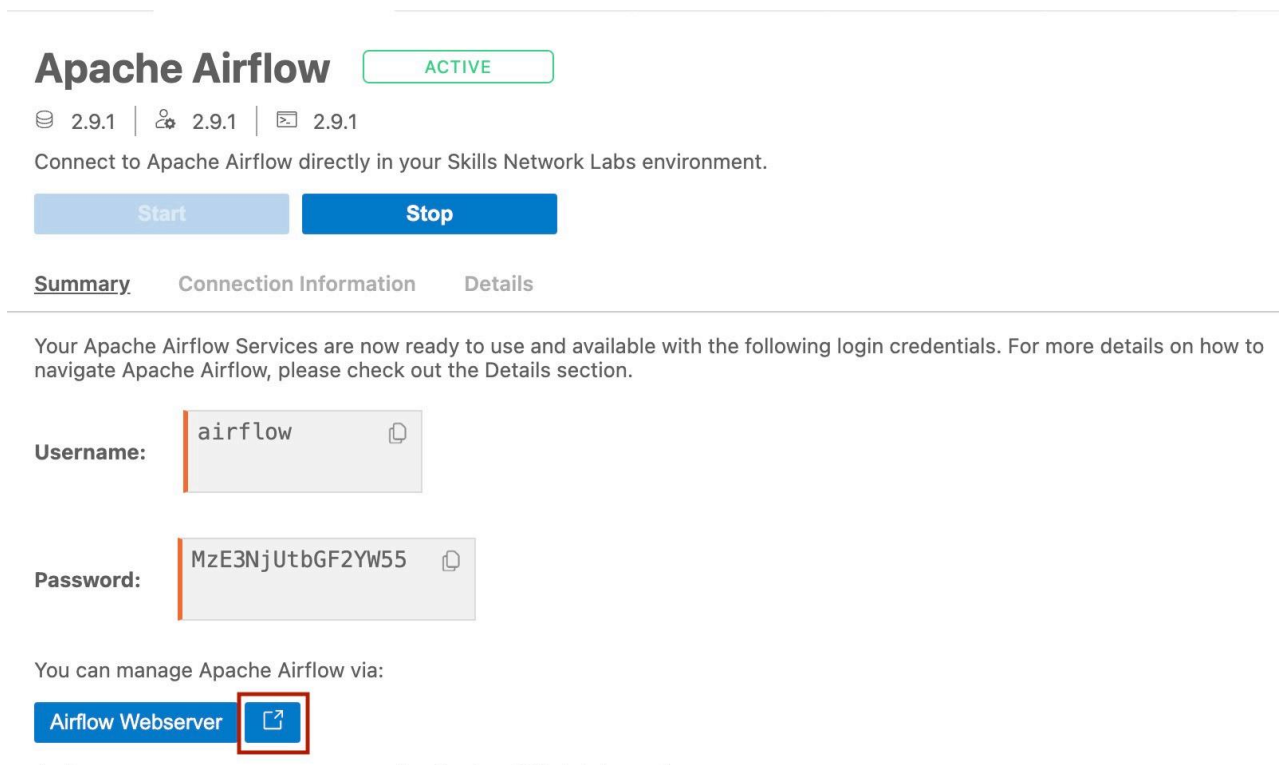


**Nota:** Por favor, seja paciente, levará alguns minutos para o Airflow iniciar.



## Exercício 2: Abra a Interface Web do Airflow

Quando o Airflow iniciar com sucesso, você deve ver uma saída semelhante à abaixo. Assim que **Apache Airflow** tiver iniciado, clique no ícone destacado para abrir a **Interface Web do Apache Airflow** em uma nova janela.



Você deve chegar a uma página que se parece com isso.



```
export AIRFLOW_HOME=/home/project/airflow
```

4. Submeter um DAG é tão simples quanto copiar o arquivo python do DAG para a pasta dags no diretório AIRFLOW\_HOME. Abra um terminal e execute o comando abaixo para submeter o DAG.

```
cp dummy_dag.py $AIRFLOW_HOME/dags
```

5. Verifique se nosso DAG foi realmente submetido. Execute o comando abaixo para listar todos os DAGs existentes.

```
airflow dags list
```

6. Verifique se dummy\_dag faz parte da saída.

```
airflow dags list | grep dummy_dag
```

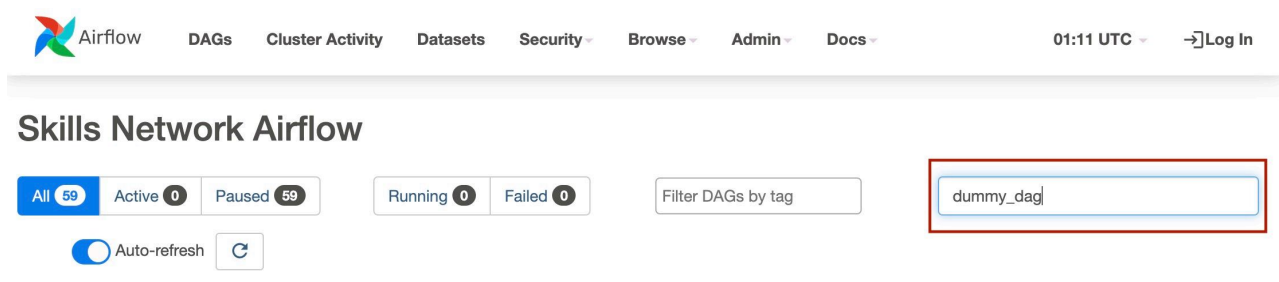
7. Execute o comando abaixo para listar todas as tarefas em dummy\_dag.

```
airflow tasks list dummy_dag
```

Você deve ver 3 tarefas na saída.


## Exercício 4: Buscar por um DAG

1. Na interface Web, identifique a caixa de texto Search DAGs conforme mostrado na imagem abaixo e digite dummy\_dag na caixa de texto e pressione enter.



Nota: Pode levar alguns minutos para que o DAG apareça aqui. Se você não vir seu DAG, por favor, aguarde um minuto e tente novamente.

2. Você deverá ver o dummy\_dag listado, conforme visto na imagem abaixo:



01:17 UTC
Log In

























[DAGs](#)
[Cluster Activity](#)
[Datasets](#)
[Security](#)
[Browse](#)
[Admin](#)
[Docs](#)

## Skills Network Airflow

All 1
Active 0
Paused 1

Running 0
Failed 0

☒ Auto-refresh
 

DAG	Owner	Runs	Schedule	Last Run	Next Run	Recent Tasks
<input checked="" type="checkbox"/> dummy_dag	Your name	   	0:01:00	2024-06-06, 00:00:00	                   	

### Exercício 5: Pausar/Despausar um DAG

1. Despause o DAG usando o botão Pausar/Despausar.

Skills Network Airflow

01:20 UTC Log In

[DAGs](#)
[Cluster Activity](#)
[Datasets](#)
[Security](#)
[Browse](#)
[Admin](#)
[Docs](#)

[All 1](#)
[Active 0](#)
[Paused 1](#)
[Running 0](#)
[Failed 0](#)

Filter DAGs by tag

dummydag

Auto-refresh

DAG	Owner	Runs	Schedule	Last Run	Next Run	Recent Tasks	Actions	Links
<div> <div>Pause/Unpause DAG</div> <div>dummydag</div> </div>	Your name	<div> <div></div> <div></div> <div></div> </div> 0:01:00		2024-06-06, 00:00:00	<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></div> </div>	...	


2. Você pode ver os seguintes detalhes nesta visualização.

- Proprietário do DAG
- Quantas vezes este DAG foi executado
- Cronograma do DAG
- Último horário de execução do DAG
- Status das tarefas recentes

[illegible]

### Exercício 6: Visão detalhada de um DAG

1. Clique no nome do DAG conforme mostrado na imagem abaixo para ver a visão detalhada do DAG.


Airflow

[DAGs](#)
[Cluster Activity](#)
[Datasets](#)
[Security](#)
[Browse](#)
[Admin](#)
[Docs](#)

01:27 UTC
[→ Log In](#)

## Skills Network Airflow


All 1

Active 0

Paused 1

Running 0

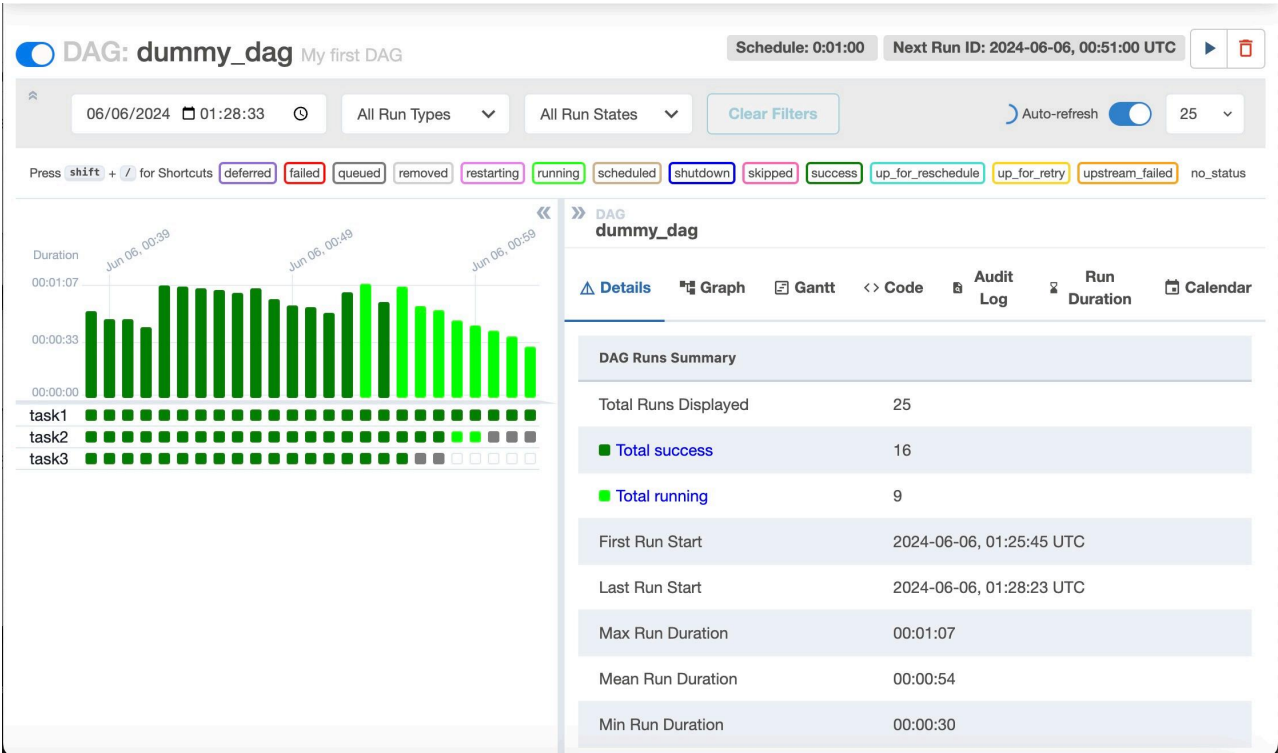
Failed 0

☒ Auto-refresh
 

DAG	Owner	Runs	Schedule	Last Run	Next Run	Recent Tasks	Actions	Links
<input checked="" type="checkbox"/> <b>dummy_dag</b>	Your name	<div> <div>40</div> <div>5</div> </div>	0:01:00	2024-06-06, 00:50:00	2024-06-06, 00:00:00	<div> <div>3</div> <div>3</div> <div>2</div> <div>7</div> </div>	<input type="button" value="▶"/> <input type="button" value="🗑️"/>	...

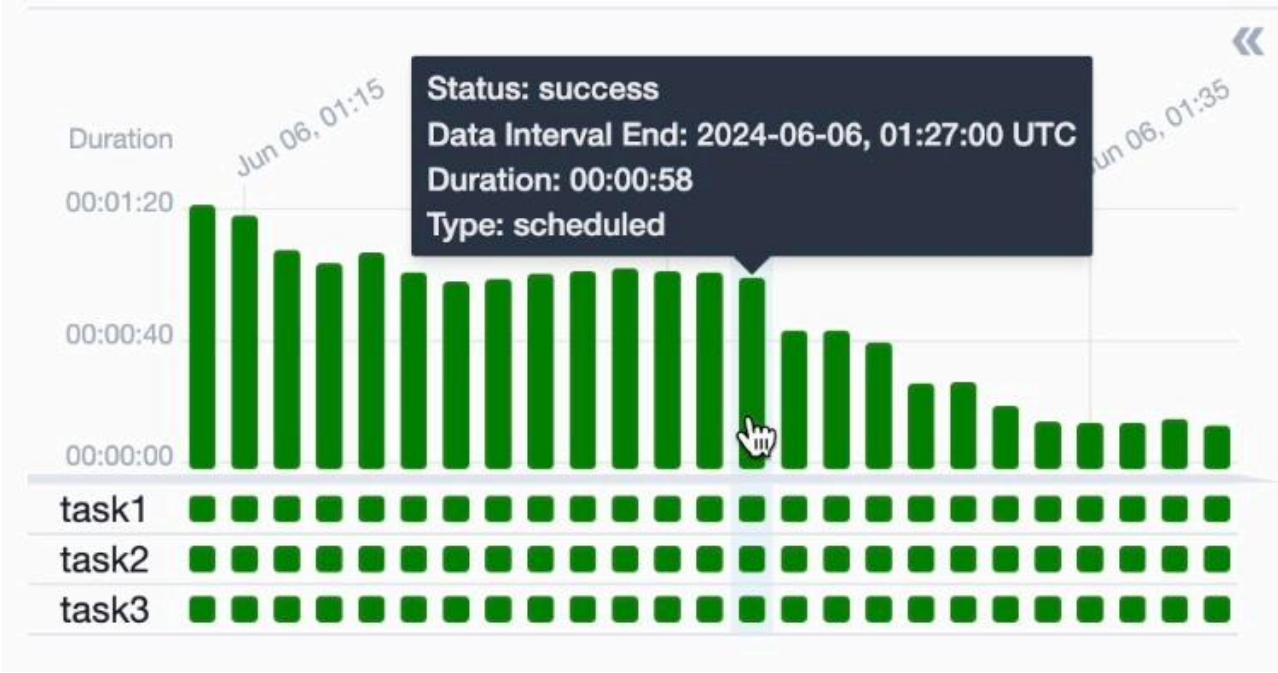
My first DAG

2. Você será direcionado para uma página de detalhes do DAG mostrando a visualização em grade padrão com as três tarefas listadas.

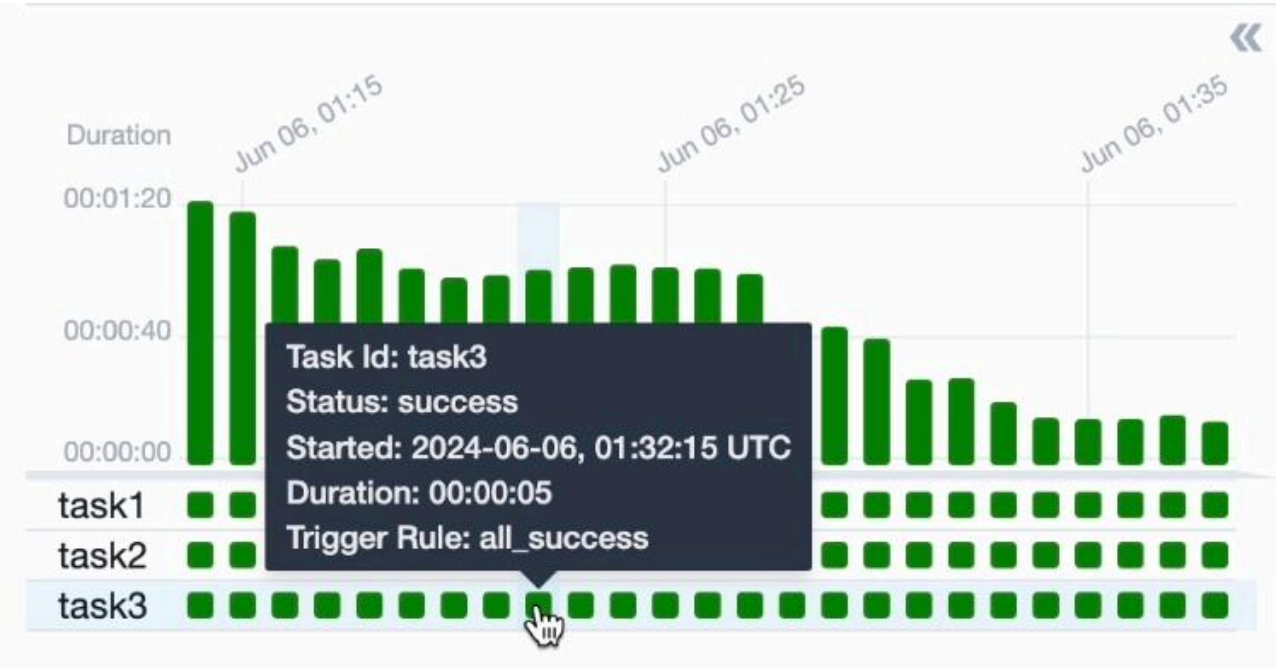


A visualização em grade mostra suas tarefas do DAG na forma de grades, como visto na imagem. Você notará que o botão Auto Refresh está ativado por padrão no canto direito.

As grades na imagem representam uma única execução do DAG e a cor indica o status da execução do DAG. Passe o mouse sobre qualquer grade para ver os detalhes.

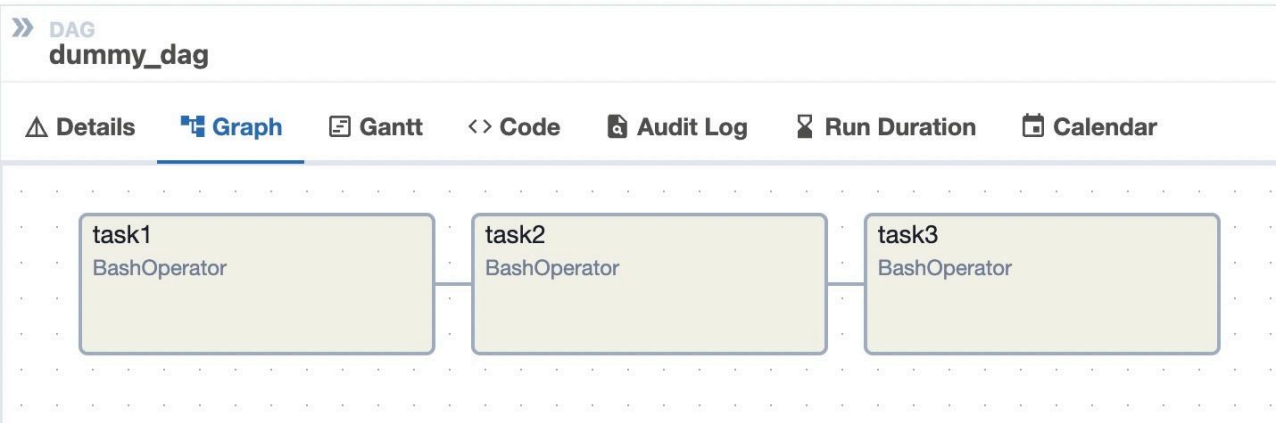


Os quadrados na imagem abaixo representam uma única tarefa dentro de uma execução do DAG e a cor indica seu status. Passe o mouse sobre qualquer quadrado para ver os detalhes da tarefa.



Exercício 7: Explorar a visualização em gráfico do DAG

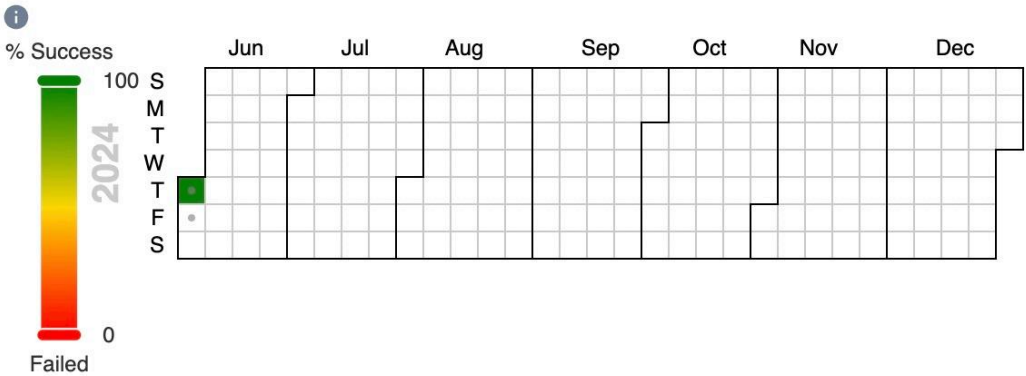
1. Clique no botão Graph view para abrir a visualização em gráfico. A visualização em gráfico mostra as tarefas na forma de um gráfico. Com a atualização automática ativada, o status de cada tarefa também é indicado com o código de cores.



Exercício 8: Visão do calendário

A visão do calendário fornece uma visão geral de todas as datas em que este DAG foi executado, juntamente com seu status representado por um código de cores.



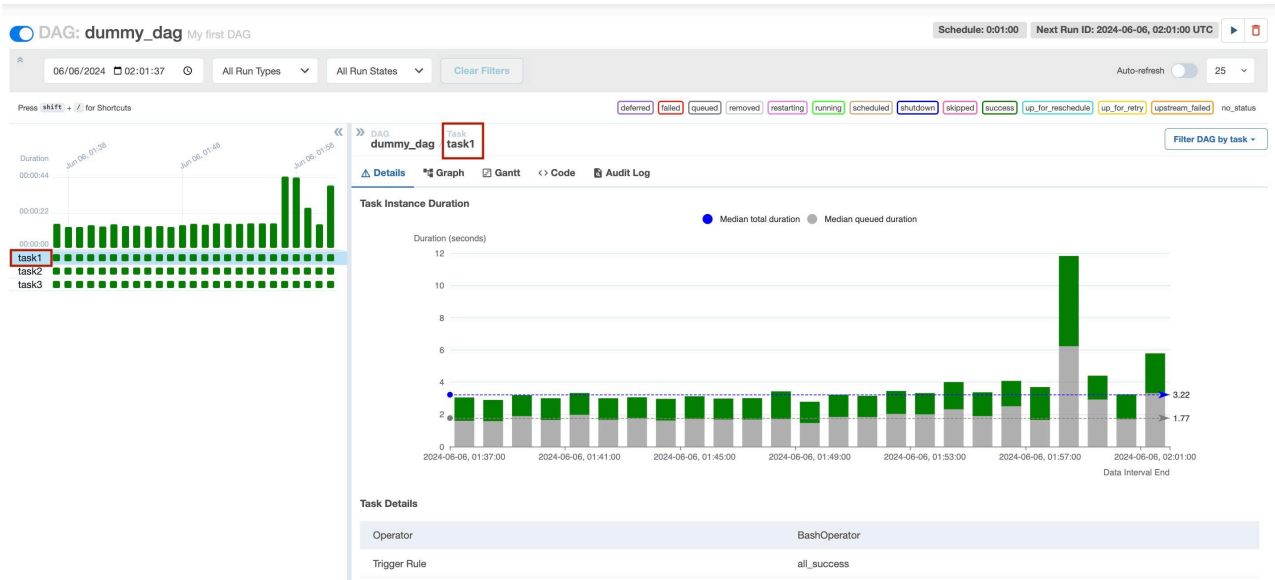


Exercício 9: Visualização de DAG e Duração de Tarefas

A duração do DAG fornece uma visão geral de quanto tempo todo o fluxo de trabalho levou.



A visualização de Duração de Tarefas oferece uma visão geral de quanto tempo cada tarefa levou para ser executada, ao longo de um período de tempo.





## Exercício 10: Visão de detalhes

A visão de detalhes fornece todas as informações do DAG conforme especificado no código do DAG.

>> DAG

dummy\_dag

Details

Graph

Gantt

<> Code

Audit Log

Run Duration

Calendar

DAG Runs Summary

Total Runs Displayed

25

Total success

25

First Run Start

2024-06-06, 01:32:20 UTC

Last Run Start

2024-06-06, 01:54:00 UTC

Max Run Duration

00:00:38

Mean Run Duration

00:00:16

Min Run Duration

00:00:13

DAG Summary

Total Tasks

3

BashOperators

3

DAG Details

Dag display name

dummy\_dag

Dag id

dummy\_dag

Description

My first DAG

Fileloc

/home/project/airflow/dags/dummy\_dag.py

Has import errors

false

Has task concurrency limits

false

Is active

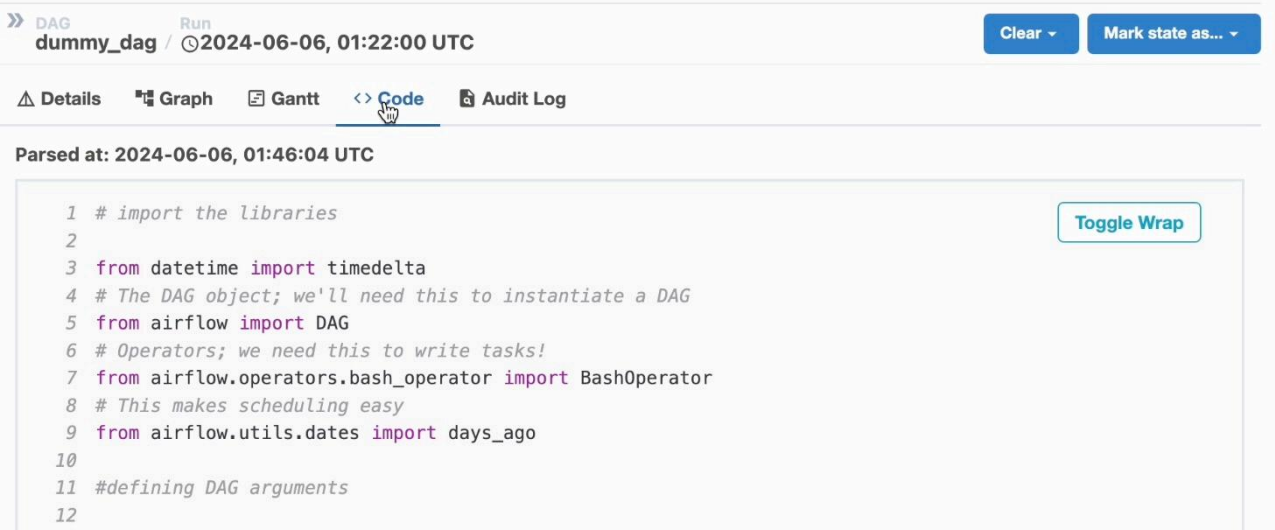
true

Is paused

false

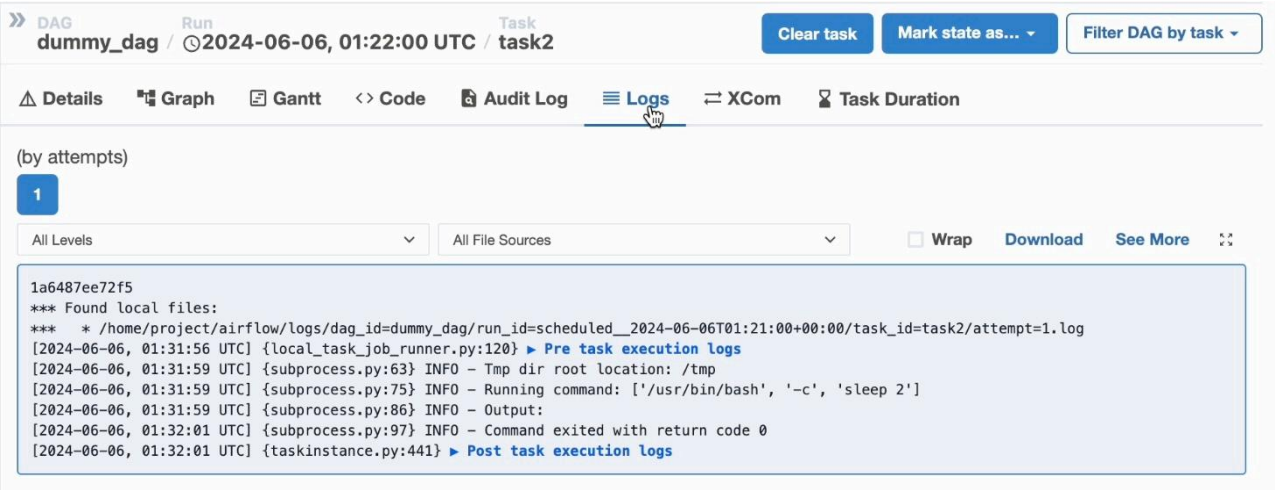
## Exercício 11: Visualização de Código

A visualização de código permite que você veja o código do DAG.



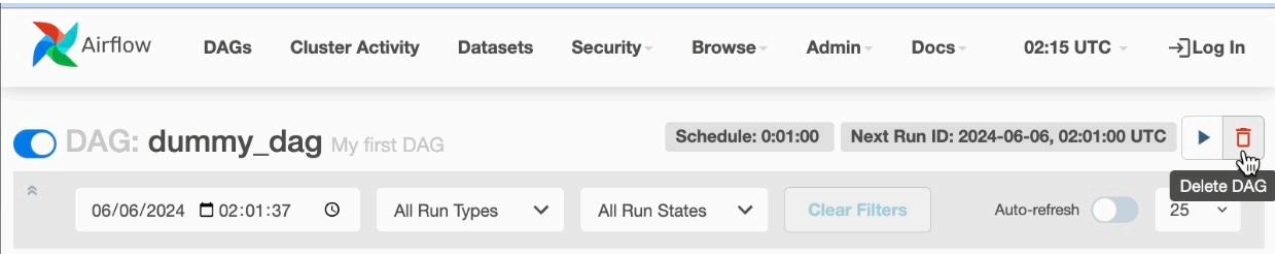
Exercício 12: Registros de tarefas

Você pode visualizar os registros de uma tarefa individual com os registros de tarefas.

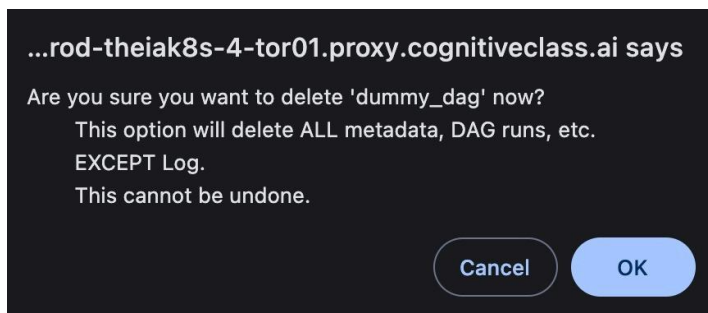


Exercício 13: Deletar um DAG

Para deletar um DAG, clique no botão de exclusão.



Você receberá um pop-up de confirmação, conforme mostrado na imagem abaixo. Clique OK para deletar o DAG.



## Exercícios de prática

1. Despause qualquer DAG existente e monitore-a.
2. Veja os detalhes de qualquer DAG existente. Veja o código da DAG. Mergulhe nos detalhes da tarefa e veja os logs de cada tarefa.

### Authors

[Lavanya T S](#)

Ramesh Sannareddy

### Other Contributors

Rav Ahuja

© IBM Corporation. Todos os direitos reservados.