

Laboratório Prático: Introdução ao Apache Airflow



Skills
Network

Tempo estimado: **20** minutos

Introdução

Neste laboratório, você terá a oportunidade de trabalhar com o Apache Airflow. Você irá utilizar tanto a interface web (UI) quanto a interface de linha de comando (CLI) para explorar as várias maneiras de usar o Airflow para gerenciamento de fluxos de trabalho.

Objetivos

Após concluir este laboratório, você será capaz de:

- Iniciar o Apache Airflow
- Abrir a UI do Airflow em um navegador
- Listar todos os DAGs
- Listar as tarefas em um DAG
- Explorar um DAG na UI

Sobre o Skills Network Cloud IDE

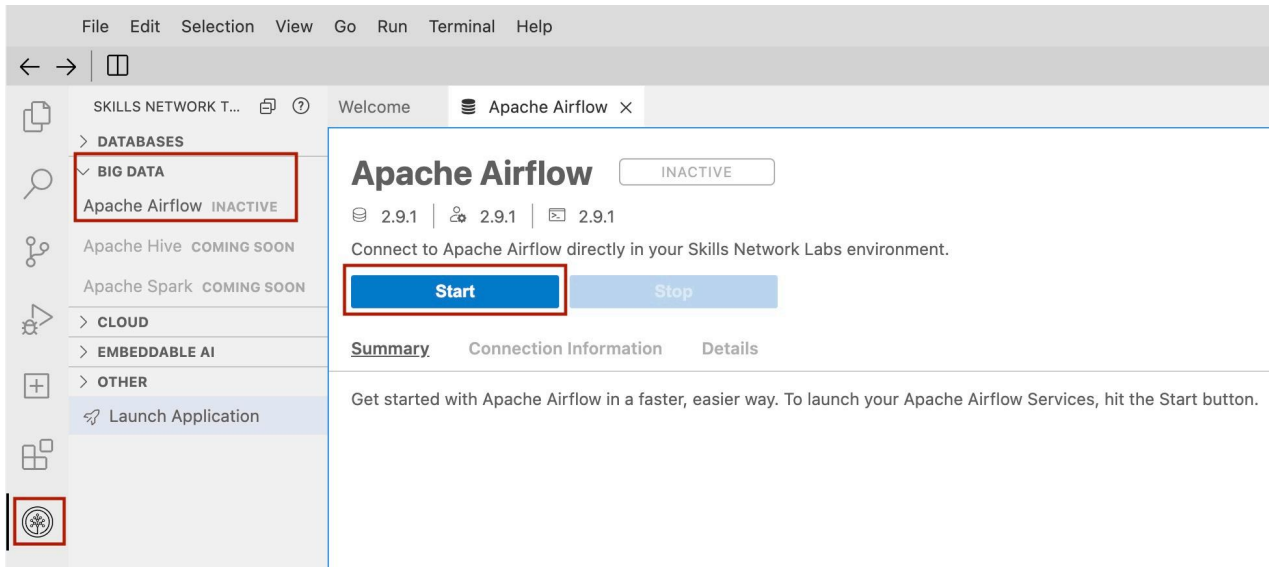
O Skills Network Cloud IDE (baseado no Theia e Docker) fornece um ambiente para laboratórios práticos relacionados a cursos e projetos. O Theia é um IDE (Ambiente de Desenvolvimento Integrado) de código aberto, que pode ser executado no desktop ou na nuvem. Para completar este laboratório, usaremos o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

Exercício 1: Iniciar Apache Airflow

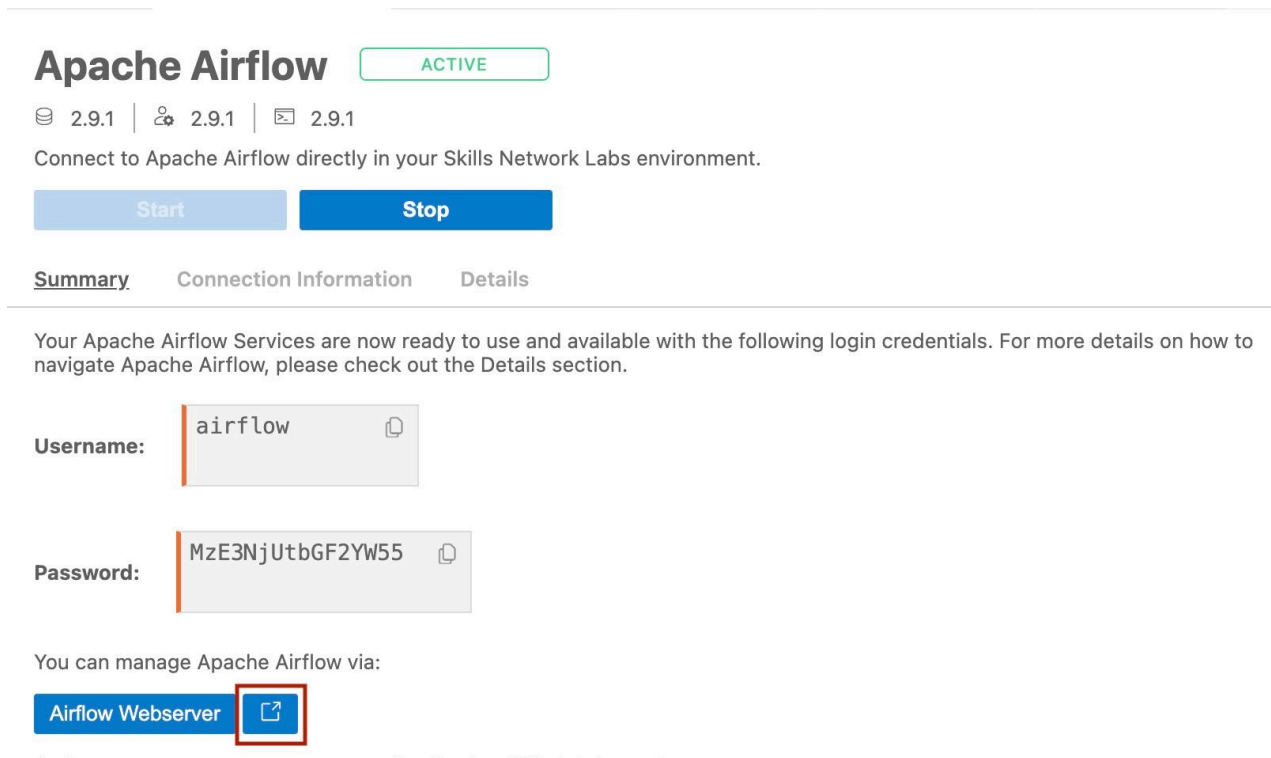
1. Clique em **Skills Network Toolbox**.
2. Na seção **BIG DATA**, clique em **Apache Airflow**.
3. Clique em **Iniciar** para iniciar o Apache Airflow.



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

Exercício 2: Abra a Interface Web do Airflow

1. Quando o Airflow iniciar com sucesso, você deverá 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.



2. Você será direcionado a uma página que exibe vários DAGs e opções associadas.

Airflow									
DAGs		Cluster Activity	Datasets	Security	Browse	Admin	Docs	03:53 UTC - Log In	
<input type="checkbox"/>	example_bash_operator	airflow	00000	00***				2024-05-29, 00:00:00	
<input type="checkbox"/>	example_branch_datetime_operator	airflow	00000	@daily				2024-05-29, 00:00:00	
<input type="checkbox"/>	example_branch_datetime_operator_2	airflow	00000	@daily				2024-05-29, 00:00:00	
<input type="checkbox"/>	example_branch_datetime_operator_3	airflow	00000	@daily				2024-05-29, 00:00:00	
<input type="checkbox"/>	example_branch_dop_operator_v3	airflow	00000	*/1****				2024-05-30, 03:51:00	
<input type="checkbox"/>	example_branch_labels	airflow	00000	@daily				2024-05-29, 00:00:00	
<input type="checkbox"/>	example_branch_operator	airflow	00000	@daily				2024-05-29, 00:00:00	
<input type="checkbox"/>	example_branch_python_operator_decorator	airflow	00000	@daily				2024-05-29, 00:00:00	

3. Você pode ativar e desativar um DAG usando o botão alternador **Ativar/Desativar**.
O botão fica cinza quando o DAG está desativado.

Airflow

Security

Browse

Admin

Docs

Skills Network Airflow

All 32

Active 0

Paused 32

DAG

example_bash_operator

example example2

airflow

example_branch_datetime_operator_2

example

airflow

Owner

O botão não fica cinza quando o DAG está em execução.

Airflow

Security

Browse

Admin

Docs

Skills Network Airflow

All 32

Active 0

Paused 32

DAG

example_bash_operator

example example2

airflow

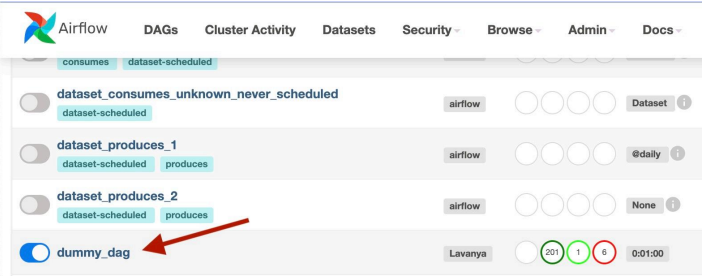
example_branch_datetime_operator_2

example

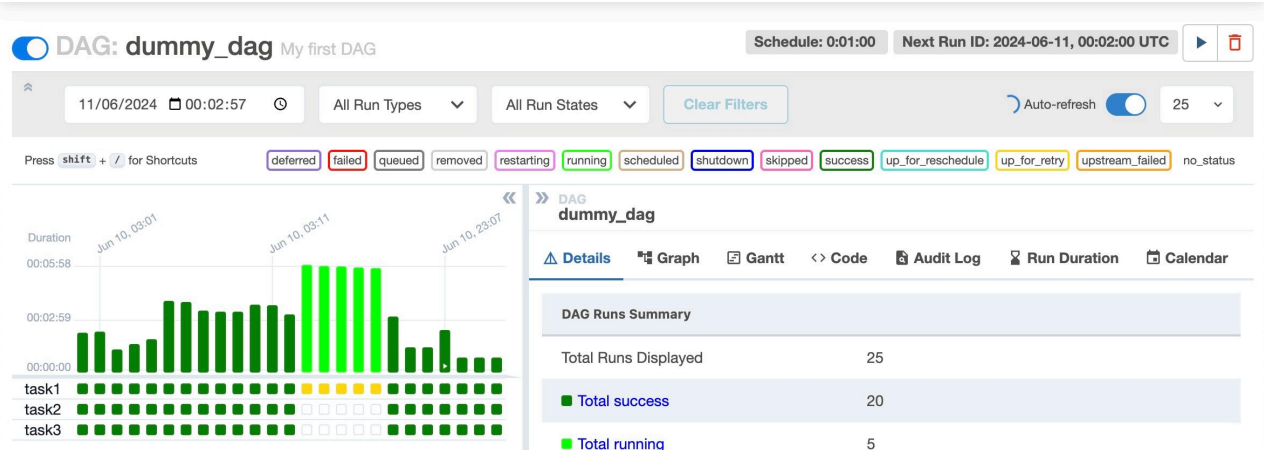
airflow

Owner

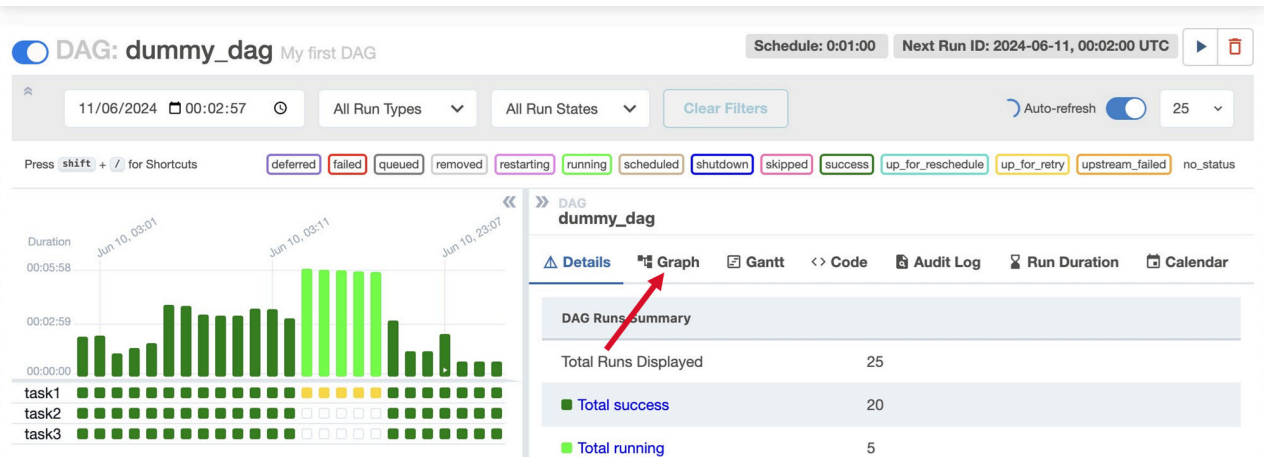
4. Clique em um DAG para explorar mais.



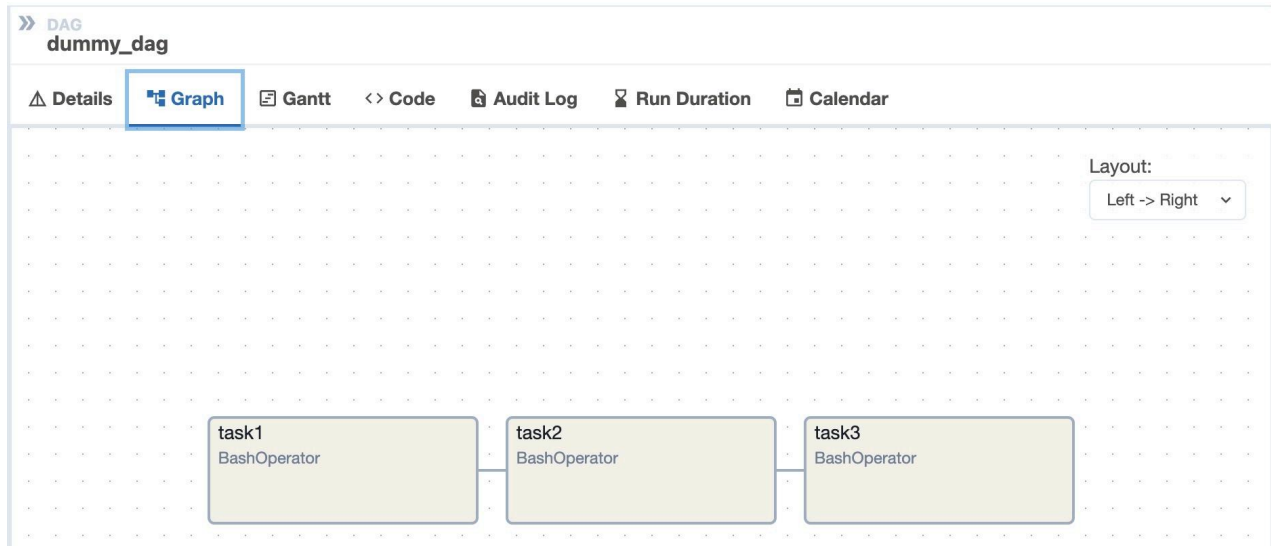
5. Você verá a visualização Grid por padrão.



6. Em seguida, clique no botão **Gráfico**.



7. Note que a visualização do gráfico do DAG será exibida.



Exercício 3: Apache Airflow CLI

Apache Airflow oferece algumas opções de linha de comando.

1. Execute o comando abaixo no terminal para listar todos os DAGs existentes.

```
airflow dags list
```

2. Execute o comando abaixo no terminal para listar todas as tarefas no DAG chamado `example_bash_operator`.

```
airflow tasks list example_bash_operator
```

Exercício Prático

1. Execute um comando para listar todas as tarefas do DAG chamado `tutorial`.
► [Clique aqui para a solução.](#)

Exercício 4: Pausar ou Retomar um DAG

1. Execute o comando abaixo no terminal para retomar um DAG chamado tutorial.

```
airflow dags unpause tutorial
```

2. Execute o comando para pausar o DAG.

```
airflow dags pause tutorial
```

Exercício Prático

1. Execute um comando para retomar o DAG chamado example_branch_operator.

▶ [Clique aqui para a solução.](#)

Exercícios de prática

1. Liste as tarefas para o DAG example_branch_labels.

▶ [Clique aqui para uma dica.](#)

▶ [Clique aqui para a solução.](#)

2. Despause o DAG example_branch_labels.

▶ [Clique aqui para uma dica.](#)

▶ [Clique aqui para a solução.](#)

3. Pause o DAG example_branch_labels.

▶ [Clique aqui para uma dica.](#)

▶ [Clique aqui para a solução.](#)

Autores

Ramesh Sannareddy

[Lavanya T S](#)

Outros Contribuintes

Rav Ahuja

© IBM Corporation. Todos os direitos reservados.