[sfeir] Institute

Google Cloud | T-WOCC-I-1.0

Workflow Orchestration with Cloud Composer

The best course to start your Data Engineering journey on Google Cloud.

1 jour / 7h

Course overview

Cloud Composer is a fully managed workflow orchestration service built on Apache Airflow. Composer enables you to create, schedule, monitor, and manage workflow pipelines that span across clouds and on-premises data centers.

In this course, you will learn about Apache Airflow and its implementation via Cloud Composer. You will learn how to provision Composer instances, create and manage Airflow DAGs on Composer, and perform tasks such as testing, debugging, and monitoring of Airflow DAGs.

Learning outcomes

- Explore Apache Airflow and Cloud Composer as workflow orchestration solutions.
- Create and manage Airflow DAGs following best practices.
- Test and debug Airflow DAGs.
- Monitor and observe Airflow DAGs on Cloud Composer.

Target audience

Data Engineers who wish to learn how to use Apache Airflow and Cloud Composer to orchestrate their data engineering workflows

Prerequisites

Completion of « Building Batch Data Pipelines on Google Cloud » or equivalent knowledge of data analytics and engineering on Google Cloud.

Course Outline

Module 01: Introduction to Cloud Composer

Topics:

- Data Engineer's need for Workflow Orchestration
- Introduction to Apache Airflow
- Cloud Composer
- Environment Setup
- Using the Composer and Airflow

Objectives:

- Explore Apache Airflow and Cloud Composer.
- Provision Cloud Composer instances.
- Explore the Airflow and Composer UIs.

Activities:

- Lab: Provisioning Cloud Composer

Module 02: Creating and managing DAGs

Topics:

- DAG structure and best practices
- Common operators
- Dependencies, trigger rules, and flow control
- Integration of Airflow and Google Cloud Services

Objectives:

- Write DAGs.

- Explore common Airflow operators.
- Manage triggers, dependencies, and flow control.
- Integrate Airflow with Google Cloud Services.

Activities:

- Lab: Assembling a Data Processing Workflow

Module 03: Advanced Airflow techniques and best practices

Topics:

- Advanced Airflow features
- Debugging DAGs
- Performance and scalability
- Security and Access Control
- Observability and monitoring

Objectives:

- Leverage advanced Airflow features.
- Debug DAGs.
- Observe and monitor your running DAGs.

Activities:

- Lab: Extending and Monitoring DAGs