

Building Batch Data Pipelines on Google Cloud

Download Whitepaper: Accelerate Your Modernization Efforts with a Cloud-Native Strategy
Get Your Free Copy Now

Course Number: GCP-BBDP

Duration: 1 days

Overview

Course Description

Data pipelines typically fall under one of the Extract and Load (EL), Extract, Load and Transform (ELT) or Extract, Transform and Load (ETL) paradigms. This course describes which paradigm should be used and when for batch data. Furthermore, this course covers several technologies on Google Cloud for data transformation including BigQuery, executing Spark on Dataproc, pipeline graphs in Cloud Data Fusion and serverless data processing with Dataflow. Learners get hands-on experience building data pipeline components on Google Cloud using Qwiklabs.

Skills Gained

- Review different methods of data loading: EL, ELT and ETL and when to use what
- Run Hadoop on Dataproc, leverage Cloud Storage, and optimize Dataproc jobs
- Build your data processing pipelines using Dataflow
- Manage data pipelines with Data Fusion and Cloud Composer

Who Can Benefit

Developers responsible for designing pipelines and architectures for data processing.

Prerequisites

- Experience with data modeling and ETL (extract, transform, load) activities.
- Experience with developing applications by using a common programming language such as Python or Java.

Audience

Course Details

Course Outline

Introduction

- In this module, we introduce the course and agenda

Introduction to Building Batch Data Pipelines

- This module reviews different methods of data loading: EL, ELT and ETL and when to use what

Executing Spark on Dataproc

- This module shows how to run Hadoop on Dataproc, how to leverage Cloud Storage, and how to optimize your Dataproc jobs.

Serverless Data Processing with Dataflow

- This module covers using Dataflow to build your data processing pipelines

Manage Data Pipelines with Cloud Data Fusion and Cloud Composer

- This module shows how to manage data pipelines with Cloud Data Fusion and Cloud Composer.

Course Summary

- Course Summary

Course Resources

- PDF links to all modules