

Streaming on Cloudera with Apache Flink

OEM: Cloudera • Duration: 2 Days (16 hrs) • Code: DOPS-246

COURSE MODULES & TOPICS

Overview

- Introduction to Apache Flink and Stream Processing
- Typical Use Cases
- Related Products

Service Deployment

- Planning Requirements
- Installation
- Flink Dashboard
- Exercise: Running a Flink Program

Flink Basics

- Execution Environment
- Flink Application Structure
- Create a Flink Project
- Build a Flink Program
- Exercise: Building a Simple Flink Program

Architecture

- Logical
- Physical
- Parallelism
- Fault Tolerance
- Data Storage

DataStream API

- DataStream API Overview
- Data Types and Serialization
- Sources and Sinks
- Transformations

- Exercise: Batch Processing Using Flink
- Exercise: Creating a Flink Streaming Application

DataStream API (continued)

- Using Kafka as a Source and Sink
- Exercise: Creating a Streaming Application Using a Kafka Source

FlinkSQL and Table API

- Streaming Concepts
- Programming Options
- Integrations
- Exercise: Using FlinkSQL and Kafka

Stateful Stream Processing

- Stateful Streaming Applications
- Checkpoints
- Event Time Processing
- Watermarks
- Windows
- Exercise: Tumbling Windows with Event Time

Cloudera SQL Stream Builder

- Overview
- Streaming SQL Console
- Data Providers
- SQL Stream Jobs
- Exercise: Using SQL Stream Builder

Monitoring

- Metrics
- Logging
- Backpressure
- Resource Configuration
- Exercise: Monitoring