

**Module Title : Cloudera DataFlow Flow Management with Apache NiFi**

**Duration : 3 days**

## Overview

One of the most critical functions of a data-driven enterprise is the ability to manage ingest and data flow across complex ecosystems. Does your team have the tools and skill sets to succeed at this?

Apache NiFi provides this capability and our three-day Cloudera Dataflow: Flow Management with Apache Nifi course delivers the foundational training you'll need to succeed with NiFi. In addition to learning NiFi's key features and concepts, participants will gain hands-on experience creating, executing, managing, and optimizing NiFi dataflows throughout a variety of scenarios.

## What Skills You Will Gain

During this course, you learn how to:

- Navigate the NiFi user interface
- Define, configure, organize, and manage dataflows
- Transform and trace data as it flows to its destination
- Track changes to dataflows with NiFi Registry
- Use the NiFi Expression Language to control dataflows
- Optimize dataflows for better performance and maintainability
- Connect dataflows with other systems, such as Apache Kafka, Apache Hive, and HDFS

## Who Should Take This Course?

This course is designed for developers, data engineers, administrators, and others with an interest in learning NiFi's innovative no-code, graphical approach to data ingest. Although programming experience is not required, basic experience with Linux is presumed, and previous exposure to big data concepts and applications is helpful.

## Course Outline

### Introduction to Cloudera Flow Management

- Overview of Cloudera Flow Management and NiFi
- The NiFi User Interface
- Demonstration: NiFi User Interface
- Exercise: Build Your First Dataflow

### Processors

- Overview of Processors
- Processor Surface Panel
- Processor Configuration
- Exercise: Start Building a Dataflow Using Processors

### Connections

- Overview of Connections
- Connection Configuration
- Connector Context Menu
- Exercise: Connect Processors in a Dataflow

### Dataflows

- Command and Control of a Dataflow
- Processor Relationships
- Back Pressure
- Prioritizers
- Labels
- Exercise: Build a More Complex Dataflow
- Exercise: Creating a Fork Using Relationships
- Exercise: Set Back Pressure Thresholds

### Process Groups

- Anatomy of a Process Group
- Input and Output Ports
- Exercise: Simplify Dataflows Using Process Groups

### FlowFile Provenance

- Data Provenance Events
- FlowFile Lineage
- Replaying a FlowFile
- Exercise: Using Data Provenance

### Dataflow Templates

- Templates Overview
- Managing Templates
- Exercise: Creating, Using, and Managing Templates

### Apache NiFi Registry

- Apache NiFi Registry Overview

- Using the Registry
- Exercise: Versioning Flows Using NiFi Registry

#### **FlowFile Attributes**

- FlowFile Attributes
- Routing on Attributes
- Exercise: Working with FlowFile Attributes

#### **NiFi Expression Language**

- NiFi Expression Language Overview
- Syntax
- Expression Language Editor
- Setting Conditional Values
- Exercise: Using the NiFi Expression Language

#### **Dataflow Optimization**

- Dataflow Optimization
- Control Rate
- Managing Compute
- Exercise: Building an Optimized Dataflow

#### **NiFi Architecture**

- NiFi Architecture Overview
- Cluster Architecture
- Heartbeats
- Managing Clusters

#### **Site-to-Site Dataflows**

- Site-to-Site Theory
- Site-to-Site Architecture
- Anatomy of a Remote Process Group
- Adding and Configuring Remote Process Groups
- Exercise: Building Site-to-Site Dataflows

#### **Cloudera Edge Management and MiNiFi**

- Overview of MiNiFi
- Example Walk-through

#### **Monitoring and Reporting**

- Monitoring from NiFi
- Overview of Reporting

- Examples of Common Reporting Tasks
- Exercise: Monitoring and Report

#### **Controller Services**

- Controller Services Overview
- Common Controller Services
- Exercise: Adding Apache Hive Controller

#### **Integrating NiFi with the Cloudera Ecosystem**

- NiFi Integration Architecture
- NiFi Ecosystem Processors
- A Closer Look at NiFi and Apache Hive
- A Closer Look at NiFi and Apache Kafka
- Exercise: Integrating Dataflows with Kafka and HDFS

#### **NiFi Security**

- NiFi Security Overview
- Securing Access to the NiFi UI
- Authentication
- Authorization
- NiFi Registry Security
- NiFi Security Summary