



CLOUDERA DATAFLOW: FLOW MANAGEMENT WITH APACHE NiFi

This course is designed for Developers, Data Engineers, Data Scientists, and Data Stewards. Though programming experience is not required, basic experience with Linux is presumed. Exposure to big data concepts and applications is helpful.

COURSE INFORMATION

This three-day hands-on training course provides the fundamental concepts and experience necessary to automate the ingest, flow, transformation, and egress of data using Apache NiFi.

Participants will create and run NiFi dataflows for a variety of scenarios. Students will gain expertise using processors, connections, and process groups, and will use NiFi Expression Language to control the flow of data from various sources to multiple destinations. Participants will monitor dataflows, examine progress of data through a dataflow, and connect dataflows to external systems such as Kafka and HDFS. After taking this course, participants will have key knowledge and expertise for configuring and managing data ingestion, movement, and transformation scenarios for the enterprise.

Through instructor-led discussion and interactive, hands-on exercises conducted by SoftSource Solutions' trainers certified by Cloudera, participants will navigate the Hadoop ecosystem, learning topics such as:

- Understand the role of Apache NiFi and MiNiFi in the Cloudera DataFlow platform
- Describe NiFi's architecture, including standalone and clustered configurations
- Use key features, including FlowFiles, processors, process groups, controllers, and connections, to define a NiFi dataflow
- Navigate, configure dataflows, and use dataflow information with the NiFi User Interface
- Trace the life of data, its origin, transformation, and destination, using data provenance
- Organize and simplify dataflows
- Manage dataflow versions using the NiFi Registry
- Use the NiFi Expression Language to control dataflows
- Implement dataflow optimization methods and available monitoring and reporting features
- Connect dataflows with other systems, such as Kafka and HDFS
- Describe aspects of NiFi security

PROTECT | OPTIMIZE | TRANSFORM

CLOUDERA
Training Partner

SoftSource Solutions Pte Ltd
<http://www.softsource.com.sg> | (65) 6746 5355 | enquiry@softsource.com.sg

©2020 SoftSource Solutions Pte Ltd. All rights reserved.
SoftSource and the SoftSource logo are trademarks or registered trademarks of SoftSource Solutions Pte Ltd.
All other trademarks are the property of their respective companies. Information is subject to change without notice.

Introduction to Apache NiFi

- Overview of Cloudera Flow Management and NiFi
- The NiFi User Interface

Processors

- Overview of Processors
- Processor Surface Panel
- Processor Configuration Panel

Connections

- Overview of Connection
- Connection Configuration Panel
- Connector Pop-Up Menu

Dataflows

- Build a DataFlow
- Command and Control of a Dataflow

Dataflows in Depth

- Processor Relationships
- Back Pressure
- Prioritizers
- Labels

Process Groups

- Anatomy of a Process Group
- Input and Output Ports

Data Provenance

- Data Provenance Events
- FlowFile Lineage
- Replaying a FlowFile

Dataflow Templates

- Templates Overview
- Managing Templates

Apache NiFi Registry

- Apache NiFi Registry Overview
- Using the Registry

FlowFile Attributes

- FlowFile Attributes Overview
- Routing on Attributes

NiFi Expression Language

- NiFi Expression Language Overview
- Syntax
- Expression Language Editor
- Setting Conditional Values

Dataflow Optimization

- NiFi Dataflow Optimization
- Control Rate
- Managing Compute
- Managing Backlog

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

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

Controller Services

- Controller Services Overview
- Common Controller Services

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

NiFi Security

- NiFi Security Overview
- Securing Access to the NiFi UI
- Authentication
- The Importance of Kerberos
- NiFi Registry Security
- NiFi Security Summary

PROTECT | OPTIMIZE | TRANSFORM**CLOUDERA**
Training Partner**SoftSource Solutions Pte Ltd**
<http://www.softsource.com.sg> | (65) 6746 5355 | enquiry@softsource.com.sg©2020 SoftSource Solutions Pte Ltd. All rights reserved.
SoftSource and the SoftSource logo are trademarks or registered trademarks of SoftSource Solutions Pte Ltd.
All other trademarks are the property of their respective companies. Information is subject to change without notice.