Faysal Aziz

Apache NiFi

2-Day Course Outline

FA-0100 · 2-Day

Software Development

Introduction

This course has been designed and curated by trainers who possess vast industry experience, having collaborated with top R&D institutions internationally. Over the next two days, we will immerse ourselves in the intricate world of NiFi, explore its capabilities, and understand how to harness its potential for your enterprise needs.

Apache NiFi is an open-source, distributed data integration and automation platform designed to automate the flow of data between software systems. It provides a web-based interface for designing, controlling, and monitoring data flows.

Learning outcomes

By the end of this training, participants will be able to:

- Understand and articulate the challenges of enterprise data flow and discern how Apache NiFi addresses these problems.
- Grasp key NiFi concepts including FlowFile, Parameter Context, Variables, Expression Language, and Funnels.
- Install and set up Apache NiFi on an Ubuntu system.
- Build NiFi data flow pipelines for real-world scenarios.
- Monitor resource usage and implement logging for NiFi pipelines.
- Integrate NiFi processes with Python and Java applications.

Prerequisites

- Basic understanding of data flow and data processing concepts.
- Familiarity with data storage and networking basics.

- A computer system running Ubuntu.
- Basic knowledge of Python and Java programming.

Training outline

1. Introduction to Apache NiFi:

1. Background and evolution of NiFi.
2. Core benefits of adopting NiFi for data flow management.
3. Architecture and foundational components of NiFi.

2. Understanding Enterprise Data Flow Challenges:

1. Overview of prevalent data flow problems in organizations.

   Solutions offered by NiFi:

   1. Scalability.
   2. Data provenance.
   3. Secure data transmission.
   4. Flexibility and extensibility.

3. Diving into NiFi Key Concepts:

   FlowFile: Delving into NiFi's core data unit.

   1. Attributes and content.
   2. Lifecycle insights and modifications.

2. Parameter Context: Making NiFi components dynamic with parameters.
3. Variables: Their roles and benefits in the flow design.

   Expression Language: Powering dynamic data flows.

   1. Basic syntax and operations.
   2. Practical exercises.

5. Funnels: Directing and controlling data flow.

4. Installing Apache NiFi on Ubuntu:

1. Pre-installation requirements and checks.
2. Step-by-step installation process.
3. Initial configuration and setup.

4. Verifying the installation.

5. Building NiFi Data Flow Pipelines:

1. Guiding principles for effective pipeline design.

   Hands-on sessions:

   1. Ingesting data from varied sources.
   2. Transforming and routing data efficiently.
   3. Combining multiple data streams.

6. Logging, Monitoring, and Resource Management:

1. Importance of vigilance in data flow operations.
2. Configuring logging mechanisms for NiFi components.
3. Techniques for overseeing resource consumption.
4. Analyzing logs for insights and potential optimizations.

7. Integration with Python and Java:

1. Overview of NiFi's integration capabilities.
2. Setting up and using NiFi processors for Python scripts.
3. Integrating Java applications with NiFi processes.
4. Hands-on: Building a simple NiFi flow integrated with Python and Java.

Upon the conclusion of this training, participants will have a thorough understanding of Apache NiFi's primary functionalities and will be proficient in utilizing its features to address contemporary data flow challenges.

© 2026 Faysal Aziz · Kuala Lumpur