Due to the nature of this material, this document refers to numerous hardware and software products by their trade names. References to other companies and their products are for informational purposes only, and all trademarks are the properties of their respective companies. It is not the intent of ProTech Professional Technical Services, Inc. to use any of these names generically.

Course Outline

Hortonworks HDF NiFi Flow Management

Course Summary

Description

This course is designed for Data Stewards or Data Flow Managers who are looking forward to automate the flow of data between systems. Topics include Introduction to NiFi, Installing and Configuring NiFi, Detail explanation of NiFi User Interface, Explanation of its components and Elements associated with each, How to Build a dataflow, NiFi Expression Language, Understanding NiFi Clustering, Data Provenance, Security around NiFi, Monitoring Tools and HDF best practices.

Topics

- Introduction to Enterprise Data Flow
- What’s new in HDF 3.0
- HDF 3.0 Architecture and Features
- HDF System Requirements
- Install and Configure HDF [NiFi]
- Describe NiFi User Interface in detail
- Describe NiFi UI Summary and History section
- Describe Anatomy of a Processor
- Describe Anatomy of a Connection
- Describe Controller Services and Reporting Tasks
- Learn How to Build a NiFi Data Flow
- Command and Control of a NiFi Data Flow
- Describe Anatomy of a Process Group
- Remote Processor Group Transmission
- NiFi Site-to-Site Communication
- Describe the Function and Purpose of the NiFi Expression Language
- Structure of a NiFi Expression
- How to Use Expression Language Functions
- Using Expression Language Editor
- Using If/Then/Else in NiFi Expression Language
- Using Attributes and Properties
- Create, Manage and Instantiate a NiFi Templates
- How to optimize an HDF Data Flow
- Define Data Provenance and Data Provenance Events
- Describe NiFi Cluster and State Management
- Describe Cluster Setup and Management via NiFi UI
- Explain the Mechanisms Available for NiFi Monitoring
- Describe How HDF Complements the Hortonworks Data Platform (HDF)
- Describe how Big Data Ingestion is possible with HDF
- Describe HDF Configuration Best Practices
- Describe the Process of Securing HDF with 2-Way-SSL
- Describe LDAP User Authentication with NiFi
- Describe Kerberos Authentication with NiFi
- Describe HDF Multi-tenancy
- Describe how File Based Authorizer in NiFi works
- Describe how Ranger Based Authorizer in NiFi works
- Describe the Architecture of Authorization Via the Ranger-NiFi Plug-in
- List the Installation Prerequisites, configure and Install Ranger
- Describe how to create Ranger policies for NiFi

Audience

This course is designed for data engineers, integration engineers, and architects who are looking to automate data flow between systems.

Prerequisites

Before taking this course, students should be familiar with programming principles and have previous experience in software development. Experience with Linux and a basic understanding of Dataflow tools would be helpful. No prior Hadoop/NiFi experience required, but is very helpful.

Duration

Three days
Hortonworks HDF NiFi Flow Management

Course Outline

I. Day 1 Objectives
A. Introduction to Enterprise Data Flow
B. What's new in HDF 3.0
C. HDF 3.0 Architecture and Features
D. HDF System Requirements
E. Install and Configure HDF [NiFi]
F. Describe NiFi User Interface in detail
G. Describe NiFi UI Summary and History section
H. Describe Anatomy of a Processor
I. Describe Anatomy of a Connection
J. Describe Controller Services and Reporting Tasks
K. Learn How to Build a NiFi Data Flow
L. Command and Control of a NiFi Data Flow
M. Describe Anatomy of a Process Group

Day 1 Labs and Demonstration
- Installing and Starting HDF with Ambari
- Demonstration: NiFi User Interface in detail
- Building a NiFi Dataflow
- Working with NiFi Process Groups

II. Day 2 Objectives
A. Anatomy of a Remote Processor Group
B. Remote Processor Group Transmission
C. NiFi Site-to-Site Communication
D. Describe the Function and Purpose of the NiFi Expression Language
E. Structure of a NiFi Expression
F. How to Use Expression Language Functions
G. Using Expression Language Editor
H. Using If/Then/Else in NiFi Expression Language
I. Using Attributes and Properties
J. Create, Manage and Instantiate a NiFi Templates
K. How to optimize an HDF Data Flow
L. Define Data Provenance and Data Provenance Events
M. Describe NiFi Cluster and State Management
N. Describe Cluster Setup and Management via NiFi UI
O. Explain the Mechanisms Available for NiFi Monitoring

Day 2 Labs and Demonstrations
- Working with Remote Processor Groups Site-to-Site
- Working with the NiFi Expression Language
- Demonstration: Working with Attributes
- Demonstration: Working with Templates
- Demonstration: Data Provenance
- Working with NiFi Clusters
- Demonstration: NiFi Notification Services
- Demonstration: NiFi Monitoring
- Advanced NiFi Monitoring

III. Day 3 Objectives
A. Describe How HDF Complements the Hortonworks Data Platform (HDF)
B. Describe how Big Data Ingestion is possible with HDF
C. Describe HDF Configuration Best Practices
D. Describe the Process of Securing HDF with 2-Way-SSL
E. Describe LDAP User Authentication with NiFi
F. Describe Kerberos Authentication with NiFi
G. Describe HDF Multi-tenancy
H. Describe how File Based Authorization in NiFi works
I. Describe how Ranger Based Authorizer in NiFi works
J. Describe the Architecture of Authorization Via the Ranger-NiFi Plugin
K. List the Installation Prerequisites, configure and Install Ranger
L. Describe how to create Ranger policies for NiFi

Day 3 Labs and Demonstrations
- Integrating HDF with HDP
- Securing HDF with 2-way SSL Using Ambari
- NiFi User Authentication with LDAP
- Installing Ranger and Configuring NiFi with Kerberos
- Working with File Based Authorization in NiFi
- Working with Ranger Based Authorization in NiFi