

[ADM-303]: HDF Operations: Hadoop Admin II

Length : 3 Days
Delivery Method : Instructor-led (Classroom)

Course Overview

This course is designed for experienced administrators who manage Hortonworks Data Platform (HDP) 2.3 clusters with Ambari. It covers upgrades, configuration, application management, and other common tasks.

Audience Profile

IT administrators and operators responsible for configuring, managing, and supporting an Hadoop 2.3 deployment in a Linux environment using Ambari.

Pre-Requisites

Attendees should have attended HDP Operations: Hadoop Administration 1 or possess equivalent knowledge and experience. Attendees should be familiar with basic HDP administration and Linux environments.

Formats

Lecture/Discussion	50%
Hands-on Labs	50%

Course Outline

Module 1: Performing a Rolling Upgrade, Configuring Heterogeneous HDFS Storage Lessons

- List the HDP Upgrade Types
- List the HDP Upgrade Path Restrictions
- Preparing Databases and HDFS
- Describe the Process Used to Register a New HDP Version
- Execute Automated Installation of and Upgrades to HDP Clusters
- Identify the Purpose and Operation of Heterogenous HDFS Storage
- Identify and Describe the HDFS Storage Types and Policies
- Identify HDFS NFS Gateway Use Cases
- Recall HDFS NFS Gateway Architecture and Operation
- Install and Configure an HDFS NFS Gateway
- Configure an HDFS NFS Gateway Client
- Summarize the Purpose and Operation of HDFS Centralized Caching
- Configure HDFS Centralized Cache
- Define and Manage Cache Pools and Cache Directives
- Summarize the Importance of File Format and Compression Algorithm Selection

AVANTUS TRAINING PTE LTD

80 Jurong East Street 21 #04-04 Devan Nair Institute Singapore 609607
Main Line: +65 6661 0888 | Fax: +65 6661 0886
Email: enquiries@AvantusTraining.com
www.AvantusTraining.com

- Summarize the Benefits and Considerations When Using Compression in Hadoop
- Describe the Administrator and Non-Administrator Roles in Managing Compression
- List and Describe Splittable Compression Formats
- Configure Default MapReduce and Tez Compression Algorithms

LABS

- Setting Up the Lab Environment
- Performing a Rolling Upgrade
- Configuring HDFS Storage Policies
- Configuring an NFS Gateway
- Configuring HDFS Centralized Cache

Module 2: Deploying Applications with Slider, Integrating Ambari with LDAP and Hive Tuning

Lessons

- Summarize the Purpose and Operation of YARN Node Labels
- Create, Add, Modify and Remove Node Labels
- Configure Queues to Access Node Label Resources
- Run Test Jobs to Confirm Node Label Behavior
- Recall the Purpose, Benefits and Components of Apache Slider
- Install and Manage an Apache Slider Application Package Using the Slider View
- Identify Characteristics of Ambari Local Versus LDAP Users and Groups
- Integrate Ambari Server with LDAP
- Configure YARN Queues, Tez and Hive Properties to Support Performance Goals
- Recall Basic Facts About Hive and the Hive Architecture
- Recall the Requirements and Benefits of Hive High Availability HA
- Summarize the Hive HA Architecture and Operation
- Configure and Test Hive HA
- Recall the Purpose, Job Types, Structure and Benefits of Apache Oozie
- Install and Configure Apache Oozie Using Ambari
- Deploy and Manage a Sample Apache Oozie Workflow

LABS

- Configuring YARN Node Labels
- Deploying Applications Using Apache Slider
- Integrating Ambari with LDAP
- Configuring Hive High Availability
- Managing Workflows Using Apache Oozie

Module 3: Apache Oozie High Availability, Introduction to Falcon, Automating Using Ambari Blueprints

Lessons

- Recall the Purpose and Architecture of Apache Oozie
- Recall the Benefits of Apache Oozie High Availability (HA)

AVANTUS TRAINING PTE LTD

80 Jurong East Street 21 #04-04 Devan Nair Institute Singapore 609607

Main Line: +65 6661 0888 | Fax: +65 6661 0886

Email: enquiries@AvantusTraining.com

www.AvantusTraining.com

- Summarize the Apache Oozie HA Architecture and Operation
- Configure Apache Oozie HA
- Understand the Challenges of Data Governance in Large, Complex Environments
- Recall the Purpose and Capabilities of Falcon
- Understand the Purpose and Configuration of Cluster, Feed and Process Entities
- Create a Cluster Entity and Set Up Mirroring Using the Falcon UI
- Summarize the Purpose and Benefits of Ambari Blueprints
- Recall the Processes Used to Deploy a Cluster Using Ambari Blueprints
- Configure Ambari Blueprints Logical Cluster Configuration Files
- Configure Ambari Blueprints Cluster Creation Configuration Files
- Configure Ambari Blueprints Host Creation Configuration Files
- Identify Ambari Blueprints Configuration Property Precedence and Best Practices

LABS

- Integrating HDF with HDP
- Configuring Apache Oozie High Availability
- Configuring Data Replication Using Apache Falcon
- Automating Cluster Provisioning Using Apache Ambari

AVANTUS TRAINING PTE LTD

80 Jurong East Street 21 #04-04 Devan Nair Institute Singapore 609607

Main Line: +65 6661 0888 | Fax: +65 6661 0886

Email: enquiries@AvantusTraining.com

www.AvantusTraining.com