

TRAINING OFFERING | ADM-303

# HORTONWORKS DATA PLATFORM (HDP®) OPERATIONS: ADMINISTRATION 2

3 DAYS    
INTERMEDIATE ADMINISTRATOR

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.

## PREREQUISITES

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.

## TARGET AUDIENCE

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

## FORMAT

50% Lecture/Discussion  
50% Hands-on Labs

## AGENDA SUMMARY

**Day 1:** Performing a Rolling Upgrade, Configuring Heterogeneous HDFS Storage

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

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

## DAY 1 OBJECTIVES

- 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
- 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

## DAY 1 LABS

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

### **About Hortonworks**

Hortonworks is a leading innovator at creating, distributing and supporting enterprise-ready open data platforms. Our mission is to manage the world's data. We have a single-minded focus on driving innovation in open source communities such as Apache Hadoop, NiFi, and Spark. Our open Connected Data Platforms power Modern Data Applications that deliver actionable intelligence from all data: data-in-motion and data-at-rest. Along with our 1600+ partners, we provide the expertise, training and services that allows our customers to unlock the transformational value of data across any line of business. We are Powering the Future of Data™.

### Contact

For further information visit [www.hortonworks.com](http://www.hortonworks.com) +1 408 675-0983  
+1 855 8-HORTON  
INTL: +44 (0) 20 3826 1405

## DAY 2 OBJECTIVES

- 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

## DAY 2 LABS

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

### **About Hortonworks**

Hortonworks is a leading innovator at creating, distributing and supporting enterprise-ready open data platforms. Our mission is to manage the world's data. We have a single-minded focus on driving innovation in open source communities such as Apache Hadoop, NiFi, and Spark. Our open Connected Data Platforms power Modern Data Applications that deliver actionable intelligence from all data: data-in-motion and data-at-rest. Along with our 1600+ partners, we provide the expertise, training and services that allows our customers to unlock the transformational value of data across any line of business. We are Powering the Future of Data™.

### Contact

For further information visit [www.hortonworks.com](http://www.hortonworks.com) +1 408 675-0983  
+1 855 8-HORTON  
INTL: +44 (0) 20 3826 1405

## DAY 3 OBJECTIVES

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

## DAY 3 LABS

- Configuring Apache Oozie High Availability
- Configuring Data Replication Using Apache Falcon
- Automating Cluster Provisioning Using Apache Ambari

Revised 10/06/2017

### **About Hortonworks**

Hortonworks is a leading innovator at creating, distributing and supporting enterprise-ready open data platforms. Our mission is to manage the world's data. We have a single-minded focus on driving innovation in open source communities such as Apache Hadoop, NiFi, and Spark. Our open Connected Data Platforms power Modern Data Applications that deliver actionable intelligence from all data: data-in-motion and data-at-rest. Along with our 1600+ partners, we provide the expertise, training and services that allows our customers to unlock the transformational value of data across any line of business. We are Powering the Future of Data™.

### Contact

For further information visit [www.hortonworks.com](http://www.hortonworks.com) +1 408 675-0983  
+1 855 8-HORTON  
INTL: +44 (0) 20 3826 1405