

## **HDP Operations: Administration Foundations (ADM-221)**

### **COURSE OVERVIEW:**

This course is intended for systems administrators who will be responsible for the design, installation, configuration, and management of the Hortonworks Data Platform (HDP). The course provides in-depth knowledge and experience in using Apache Ambari as the operational management platform for HDP. This course presumes no prior knowledge or experience with Hadoop.

### **WHO WILL BENEFIT FROM THIS COURSE?**

Linux administrators and system operators responsible for installing, configuring and managing an HDP cluster.

### **PREREQUISITES:**

To fully benefit from this course, students should have the following prerequisite skills and knowledge:

Students must have experience working in a Linux environment with standard Linux system commands. Students should be able to read and execute basic Linux shell scripts. Basic knowledge of SQL statements is recommended, but not a requirement. In addition, it is recommended for students to have some operational experience in data center practices, such as change management, release management, incident management, and problem management.

### **COURSE OBJECTIVES:**

After completion of this course, students will be able to...

- Day 1: Introduction to Big Data, Hadoop and the Hortonworks Data Platform
- Day 2: Managing HDFS Storage, Rack Awareness, HDFS Snapshots and HDFS Centralized Cache
- Day 3: Introduction to YARN
- Day 4: High Availability with HDP, Deploying HDP with Blueprints, and the HDP Upgrade Process

### **COURSE OUTLINE:**

#### **DAY 1 OBJECTIVES**

- Describe Apache Hadoop
- Summarize the Purpose of the Hortonworks Data Platform Software Frameworks
- List Hadoop Cluster Management Choices
- Describe Apache Ambari
- Identify Hadoop Cluster Deployment Options
- Plan for a Hadoop Cluster Deployment
- Perform an Interactive HDP Installation using Apache Ambari
- Install Apache Ambari
- Describe the Differences Between Hadoop Users, Hadoop Service Owners, and Apache Ambari Users
- Manage Users, Groups and Permissions
- Identify Hadoop Configuration Files
- Summarize Operations of the Web UI Tool

- Manage Hadoop Service Configuration Properties Using the Apache Ambari Web UI
- Describe the Hadoop Distributed File System (HDFS)
- Perform HDFS Shell Operations
- Use WebHDFS
- Protect Data Using HDFS Access Control Lists (ACLs)

#### **DAY 1 LABS**

- Setting Up the Environment
- Installing HDP
- Managing Ambari Users and Groups
- Managing Hadoop Services
- Using HDFS Storage
- Using WebHDFS
- Using HDFS Access Control Lists

#### **DAY 2 OBJECTIVES**

- Describe HDFS Architecture and Operation
- Manage HDFS using Ambari Web, NameNode and DataNode UIs
- Manage HDFS using Command-line Tools
- Summarize the Purpose and Benefits of Rack Awareness
- Configure Rack Awareness
- Summarize Hadoop Backup Considerations
- Enable and Manage HDFS Snapshots
- Copy Data Using DistCP
- Use Snapshots and DistCP Together
- Identify the Purpose and Operation of Heterogeneous HDFS Storage
- Summarize the Purpose and Operation of HDFS Centralized Caching
- Configure HDFS Centralized Cache
- Define and Manage Cache Pools and Cache Directives
- 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

#### **DAY 2 LABS**

- Managing HDFS Storage
- Managing HDFS Quotas
- Configuring Rack Awareness
- Managing HDFS Snapshots
- Using DistCP
- Configuring HDFS Storage Policies
- Configuring HDFS Centralized Cache

- Configuring an NFS Gateway
- DAY 3 OBJECTIVES
- Describe YARN Resource Management
- Summarize YARN Architecture and Operation
- Identify and Use YARN Management Options
- Summarize YARN Response to Component Failure
- Understand the Basics of Running Simple YARN Applications
- Summarize the Purpose and Operation of the YARN Capacity Scheduler
- Configure and Manage YARN Queues
- Control Access to YARN Queues
- Summarize the Purpose and Operation of YARN Node Labels
- Describe the Process used to Create Node Labels
- Describe the Process Used to Add, Modify and Remove Node Labels
- Configure Queues to Access Node Label Resources
- Run Test Jobs to Confirm Node Label Behavior

### DAY 3 LABS

- Managing YARN Using Ambari
- Managing YARN Using CLI
- Running Sample YARN Applications
- Setting Up for Capacity Scheduler
- Managing YARN Containers and Queues
- Managing YARN ACLs and User Limits
- Working with YARN Node Labels
- DAY 4 OBJECTIVES
- Summarize the Purpose of NameNode HA
- Configure NameNode HA Using Ambari
- Summarize the Purpose of ResourceManager HA
- Configure ResourceManager HA using Apache Ambari
- Identify Reasons to Add, Replace and Delete Worker Nodes
- Demonstrate How to Add a Worker Node
- Configure and Run the HDFS Balancer
- Decommission and Re-commission a Worker Node
- Describe the Process of Moving a Master Component
- Summarize the Purpose and Operation of Apache Ambari Metrics
- Describe the Features and Benefits of the Apache Ambari Dashboard
- Summarize the Purpose and Benefits of Apache Ambari Blueprints
- Recall the Process Used to Deploy a Cluster Using Ambari Blueprints
- Recall the Definition of an HDP Stack and Interpret its Version Number
- View the Current Stack and Identify Compatible Apache Ambari Software Versions
- Recall the Types of Methods and Upgrades Available in HDP
- Describe the Upgrade Process, Restrictions and Pre-Upgrade Checklist
- Perform an Upgrade Using the Apache Ambari Web UI

**DAY 4 LABS**

- Configuring NameNode HA
- Configuring Resource Manager HA
- Adding, Decommissioning and Re-commissioning a Worker Node
- Configuring Ambari Alerts
- Deploying an HDP Cluster Using Ambari Blueprints
- Performing an HDP Upgrade - Express

**SUNSET LEARNING INSTITUTE (SLI) DIFFERENTIATORS:**

Sunset Learning Institute (SLI) has been an innovative leader in developing and delivering authorized technical training since 1996. Our goal is to help our customers optimize their cloud technology investments by providing convenient, high quality technical training that our customers can rely on. We empower students to master their desired technologies for their unique environments.

What sets SLI apart is not only our immense selection of trainings options, but our convenient and consistent delivery system. No matter how complex your environment is or where you are located, SLI is sure to have a training solution that you can count on!

**Premiere World Class Instruction Team**

- All SLI instructors have a four-year technical degree, instructor level certifications and field consulting work experience.
- Sunset Learning has won numerous Instructor Excellence and Instructor Quality Distinction awards since 2012

**Enhanced Learning Experience**

- The goal of our instructors during class is ensure students understand the material, guide them through our labs and encourage questions and interactive discussions.

**Convenient and Reliable Training Experience**

- You have the option to attend classes at any of our established training facilities or from the convenience of your home or office with the use of our HD-ILT network (High Definition Instructor Led Training)
- All Sunset Learning Institute classes are guaranteed to run – you can count on us to deliver the training you need when you need it!

**Outstanding Customer Service**

- Dedicated account manager to suggest the optimal learning path for you and your team
- Enthusiastic Student Services team available to answer any questions and ensure a quality training experience