

## HDP Operations: Install and Manage with Apache Ambari

### Overview

This course is designed for administrators who will be managing the Hortonworks Data Platform (HDP). It covers installation, configuration, maintenance, security and performance topics.

### Duration

4 days

### Target Audience

IT administrators and operators responsible for installing, configuring and supporting an Apache Hadoop 2.0 deployment in a Linux environment.

### Course Objectives

- Describe various tools and frameworks in the Hadoop 2.0 ecosystem
- Describe the Hadoop Distributed File System (HDFS) architecture
- Install and configure an HDP 2.0 cluster
- Use Ambari to monitor and manage a cluster
- Describe how files are written to and stored in HDFS
- Perform a file system check using command line and browser-based tools
- Configure the replication factor of a file
- Mount HDFS to a local filesystem using the NFS Gateway
- Deploy and configure YARN on a cluster
- Configure and troubleshoot MapReduce jobs
- Describe how YARN jobs are scheduled
- Configure the capacity and fairschedulers of the ResourceManager
- Use WebHDFS to access a cluster over HTTP
- Configure a Hiveserver
- Describe how Hive tables are created and populated
- Use Sqoop to transfer data between Hadoop and a relational database
- Use Flume to ingest streaming data into HDFS
- Deploy and run an Oozie workflow
- Commission and decommission worker nodes
- Configure a cluster to be rack-aware
- Implement and configure NameNode HA
- Secure a Hadoop cluster

### Hands-On Labs

- Install HDP 2.0 using Ambari
- Add a new node to the cluster
- Stop and start HDP services
- Use HDFS commands
- Verify data with block scanner and fsck
- Mount HDFS to a local file system
- Troubleshoot a MapReduce job
- Configure the capacity scheduler
- Use distcp to copy data from a remote cluster
- Use WebHDFS
- Use Hive tables
- Use Sqoop to transfer data
- Install and test Flume
- Run an Oozie workflow
- Commission and decommission a worker node
- Use HDFS snapshots
- Configure rack awareness
- Implement NameNode HA
- Secure an HDP cluster

### Prerequisites

Attendees should be familiar with Hadoop and Linux environments.

### Format

50% Lecture/Discussion

50% Hands-on Labs

### Certification

Hortonworks offers a comprehensive certification program that identifies you as an expert in Apache Hadoop. Visit [hortonworks.com/training/certification](http://hortonworks.com/training/certification) for more information.

### Hortonworks University

Hortonworks University is your expert source for Apache Hadoop training and certification. Public and private on-site courses are available for developers, administrators, data analysts and other IT professionals involved in implementing big data solutions. Classes combine presentation material with industry-leading hands-on labs that fully prepare students for real-world Hadoop scenarios.

#### About Hortonworks

Hortonworks develops, distributes and supports the only 100 percent open source distribution of Apache Hadoop explicitly architected, built and tested for enterprise-grade deployments.

**US:** 1.855.846.7866

**International:** 1.408.916.4121  
[www.hortonworks.com](http://www.hortonworks.com)

5470 Great American Parkway  
Santa Clara, CA 95054 USA