

# Cloudera Administrator Training for Apache Hadoop

GK3901



**Delivery Type:** Classroom

**Duration:** 3 days

## Overview

Learn to deploy, configure, and manage Cloudera's Apache Hadoop implementation and HDFS.

In this hands-on course, you will be introduced to the basics of Hadoop, Hadoop Distributed File System (HDFS), MapReduce, Hive, Pig, and HBase. You will cover core administration skills, such as cluster deployment, job management, and ongoing Hadoop maintenance and monitoring, as you gain the expertise to support your environments in day-to-day activities.

This course covers concepts addressed on the Cloudera Certified Administrator for Apache Hadoop (CCA) exam and includes a CCA exam voucher you'll receive at the end of class.

## Pre-Requisites

- ✓ Basic level of Linux system administration experience
- ✓ Prior knowledge of Apache Hadoop is not required

## Objectives

- ✓ HDFS and MapReduce
- ✓ Optimal hardware configurations for Hadoop clusters
- ✓ Network considerations to take into account when building out your cluster
- ✓ Configure Hadoop options for best cluster performance
- ✓ Configure the FairScheduler to provide service-level agreements for multiple users of a cluster
- ✓ Maintain and monitor your cluster
- ✓ Load data into the cluster from dynamically generated files using Flume and from relational database management systems using Sqoop
- ✓ System administration issues with other Hadoop projects such as Hive, Pig, and HBase

## Target Audience

System administrators looking to understand all of the steps necessary to operate and manage Apache Hadoop clusters.

## Certification

- ✓ This course is part of the following programs or tracks:
- ✓ CCAH: Cloudera Certified Administrator for Apache Hadoop (CDH3)

## Follow on Courses

- ✓ Cloudera Training for Apache HBase
- ✓ Cloudera Training for Apache Hive and Pig