HDP Cluster Administration

Overview
This course is designed for administrators who will be responsible for managing HDP 2.x (Hortonworks Data Platform). It covers extensive material for all cluster deployment, maintenance and management tasks.

Audience
All administrators responsible for supporting HDP 2.x in a Linux environment.

Duration
40 hours: Lecture Discussions
40 hours: Hands-on Labs (Homework)

Prerequisites
Students should be familiar with Linux and Hadoop environment.

Certification
This prepares you for HDPCA certification. Hortonworks offers a comprehensive certification program that identifies you as an expert in Apache Hadoop. For more information visit hortonworks.com/training/certification.

Materials
8hr Prerequisite Lecture on Linux & Hadoop.
VM for running the hands-on Lab.
Lecture and Lab materials (> 1000 pages PDF)

Class Format
Web based Instructor Led course on weekends.
4 hours of training Saturday and Sunday.

Course Objectives
Chapter 1: Big Data, Hadoop, and HDP
Chapter 2: Managing Users and Groups
Chapter 3: Managing Hadoop Services
Chapter 4: Using HDFS Storage
Chapter 5: Managing HDFS Storage
Chapter 6: YARN Resource Management
Chapter 7: YARN Data Processing
Chapter 8: Adding and Removing Nodes
Chapter 9: YARN Capacity Scheduler
Chapter 10: Rack Awareness
Chapter 11: HDFS YARN HA
Chapter 12: Monitoring Clusters
Chapter 13: Protecting a Cluster with Backups
Chapter 14: Installing HDP

Hands-on Lab
Lab 1: VM Setup Guide and Cluster Installation
Lab 2: Managing Users and Groups
Lab 3: Managing Hadoop Services
Lab 4: Using HDFS Storage
Lab 5: Using WebHDFS
Lab 6: Using HDFS ACLs
Lab 7: Managing Hadoop Storage
Lab 8: Managing HDFS Quotas
Lab 9: Managing YARN Using Ambari
Lab 10: Managing YARN Using CLI
Lab 11: Running YARN Applications
Lab 12: Adding and Removing Worker Nodes
Lab 13: Prep for Capacity Scheduler Labs
Lab 14: YARN User Limits and ACLs
Lab 15: Containers and Queues
Lab 16: Configuring Rack Awareness
Lab 17: Configuring NameNode HA
Lab 18: Configuring ResourceManager HA
Lab 19: Configuring Ambari Alerts
Lab 20: Managing HDFS Snapshots
Lab 21: Using DistCp
Lab 22: Installing HDP

Day to Day Operations (Reference Material)
Chapter 1: Hortonworks Data Platform
Chapter 2: HDFS Architecture
Chapter 3: Installation Prerequisites and Planning (Admin)
Chapter 4: Configuring Hadoop
Chapter 5: Ensuring Data Integrity
Chapter 6: HDFS NFS Gateway
Chapter 7: Yarn and MapReduce (Admin)
Chapter 8: Job Scheduling
Chapter 9: Enterprise Data Movement
Chapter 10: HDFS Web Services
Chapter 11: Hive Administration
Chapter 12: Transferring Data with Sqoop
Chapter 13: Flume
Chapter 14: Oozie (Admin)
Chapter 15: Monitoring HDP Services
Chapter 16: Commissioning and Decommissioning Nodes
Chapter 17: Backup and Recovery
Chapter 18: Rack Awareness
Chapter 20: Securing HDP