



# HDP Operations: Hadoop Fast Track Training for Beginners

## Training Calendar

Date	Training Time	Location
10 June 2019	5 Days	Bilginç IT Academy
10 June 2019	5 Days	Bilginç IT Academy

## Training Details

Training Time	:	5 Days
Capacity	:	12
Prerequisites	:	Attendees, should be familiar with server or platform software concepts and have a basic understanding of system administration.

## About Training

### About Training

This training is designed for primarily for systems administrators and platform architects who need to understand HDP cluster capabilities, and manage HDP clusters. Topics include: Understanding HDF capabilities, Apache Hadoop, Apache YARN, HDFS, and other Hadoop ecosystem components. Students will understand how to administer, manage, and monitor HDP clusters.

### What You'll Learn

- Introduction to Hadoop and Ambari
- Managing HDFS, YARN Architecture and Management

- The YARN Capacity Scheduler, High Availability, Monitoring and Backups
- Advanced HDFS & YARN Services
- Additional HDP Components and Tuning

## Who Should Attend

This course is for students who range from having an understanding of server software concepts, to system administrators and platform architects who plan on administering HDP clusters.

## Outline

### An Introduction to Apache Hadoop and HDFS

- Big Data, Hadoop and the Hortonworks Data Platform
- Installing the Hortonworks Data Platform
- Using HDFS Storage
- Managing Apache Ambari Users and Groups
- Managing Hadoop Services

### LABS

- Setting Up the Lab Environment
- Installing HDP
- Managing Apache Ambari Users and Groups
- Managing Hadoop Services

### Working with HDFS

- Using HDFS Storage
- Managing HDFS Storage
- Adding, Deleting, and Replacing Worker Nodes
- Configuring Rack Awareness

### LABS

- Using Hadoop Storage
- Using WebHDFS
- Using HDFS Access Control Lists
- Managing Hadoop Storage
- Managing HDFS Quotas
- Adding, Decommissioning, and Recommissioning Worker Nodes
- Configuring Rack Awareness

### Working with Apache YARN

- YARN Resource Management
- YARN Applications
- YARN Capacity Scheduler

### 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
- YARN Node Labels

### **High Availability, Backups and Configuring Centralized Cache**

- HDFS and YARN High Availability
  - Monitoring a Cluster
  - Protecting a Cluster with Backups
  - Configuring Heterogenous HDFS Storage
  - Managing the HDFS NFS Gateway
  - Configuring HDFS Centralized Cache

### **LABS**

- Configuring NameNode High Availability
- Configuring ResourceManager High Availability
- Managing Apache Ambari Alerts
- Managing HDFS Snapshots
- Using DistCP
- Configuring HDFS Storage Policies
- Configuring an NFS Gateway
- Configuring HDFS Centralized Cache

### **Performing a Rolling Upgrade**

- Apache Hive Tuning
- Managing Workflows Using Apache Oozie
- Integrating Ambari with LDAP
- Automating Cluster Provisioning Using Ambari Blueprints
- Performing an HDP Rolling Upgrade

### **LABS**

- Configuring Apache Hive High Availability
- Managing Workflows Using Apache Oozie
- Integrating Apache Ambari with AD/LDAP
- Automating Cluster Provisioning using Apache Ambari
- Performing an HDP Upgrade