

## HDP Overview: Apache Hadoop Essentials

### Overview

This course provides a technical overview of Apache Hadoop. It includes high-level information about concepts, architecture, operation, and uses of the Hortonworks Data Platform (HDP) and the Hadoop ecosystem. The course provides an *optional* primer for those who plan to attend a hands-on, instructor-led course

### Course Objectives

- Describe what makes data “Big Data”
- List data types stored and analyzed in Hadoop
- Describe how Big Data and Hadoop fit into your current infrastructure and environment
- Describe fundamentals of:
  - the Hadoop Distributed File System (HDFS)
  - YARN
  - MapReduce
  - Hadoop frameworks: (Pig, Hive, HCatalog, Storm, Solr, Spark, HBase, Oozie, Ambari, ZooKeeper, Sqoop, Flume, and Falcon)
  - Recognize use cases for Hadoop
  - Describe the business value of Hadoop
  - Describe new technologies like Tez and the Knox Gateway

### Hands-On Labs

- There are no labs for this course.

### Duration

8 hours – on line only

### Target Audience

Data architects, data integration architects, managers, C-level executives, decision makers, technical infrastructure team, and Hadoop administrators or developers who want to understand the fundamentals of Big Data and the Hadoop ecosystem.

### Prerequisites

No previous Hadoop or programming knowledge is required. Students will need browser access to the Internet.

### Format

- 100% self-paced, online exploration (for employees, partners or support subscription customers)
- or
- 100% instructor led discussion

### 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