HDP Analyst: Apache HBase Essentials

Overview
This course is designed for big data analysts who want to use the HBase NoSQL database which runs on top of HDFS to provide real-time read/write access to sparse datasets. Topics include HBase architecture, services, installation and schema design.

Course Objectives
- How HBase integrates with Hadoop and HDFS
- Architectural components and core concepts of HBase
- HBase functionality
- Installing and configuring HBase
- HBase schema design
- Importing and exporting data
- Backup and recovery
- Monitoring and managing HBase
- How Apache Phoenix works with HBase
- How HBase integrates with Apache ZooKeeper
- HBase services and data operations
- Optimizing HBase Access

Hands-On Labs
- Using Hadoop and MapReduce
- Using HBase
- Importing Data from MySQL to HBase
- Using Apache ZooKeeper
- Examining Configuration Files
- Using Backup and Snapshot
- HBase Shell Operations
- Creating Tables with Multiple Column Families
- Exploring HBase Schema
- Blocksize and Bloom filters
- Exporting Data
- Using a Java Data Access Object Application to Interact with HBase

Duration
2 days

Target Audience
Architects, software developers, and analysts responsible for implementing non-SQL databases in order to handle sparse data sets commonly found in big data use cases.

Prerequisites
Students must have basic familiarity with data management systems. Familiarity with Hadoop or databases is helpful but not required. Students new to Hadoop are encouraged to attend the *HDP Overview: Apache Hadoop Essentials* course.

Format
35% Lecture/Discussion
65% Hands-on Labs

Certification
Hortonworks offers a comprehensive certification program that identifies you as an expert in Apache Hadoop. Visit *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

5470 Great America Parkway
Santa Clara, CA 95054 USA