Hortonworks HDP Analyst Apache HBase Essentials

Course Summary

Description

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.

Objectives

By the end of this course, students will be able to understand:

- 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

Topics

- 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

Audience

This course is designed for 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

Before taking this course, 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.

Duration

Two days