

Hortonworks HDP Operations Apache HBase Advanced Management

Course Summary

Description

This course is designed for administrators who will be installing, configuring, and managing HBase clusters. It covers installation with Ambari, configuration, security, and troubleshooting HBase implementations. The course includes an end-of-course project in which students work together to design and implement an HBase schema.

Objectives

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

- Hadoop Primer
- Hadoop, Hortonworks, and Big Data
- HDFS and YARN
- Discussion: Running Applications in the Cloud
- Apache HBase Overview
- Provisioning the Cluster
- Using the HBase Shell
- Ingesting Data
- Operational Management
- Backup and Recovery
- Security
- Monitoring HBase and Diagnosing Problems
- Maintenance
- Troubleshooting

Topics

- Installing and Configuring HBase with Ambari
- Manually Installing HBase (Optional)
- Using Shell Commands
- Ingesting Data using ImportTSV
- Enabling HBase High Availability
- Viewing Log Files
- Configuring and Enabling Snapshots
- Configuring Cluster Replication
- Enabling Authentication and Authorization
- Diagnosing and Resolving Hot Spotting
- Region Splitting
- Monitoring JVM Garbage Collection
- End of Course Project: Designing an HBase Schema

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 take the HDP Overview: Apache Hadoop Essentials course.

Duration

Four days