



Hadoop Administration

Course description

This training course is designed for Administrator who wants to better understand how to create Apache Hadoop solutions. This 35 Hours provides Java programmers the necessary training for creating enterprise solutions using Apache Hadoop. It consists of a prudent combination of interactive lecture and extensive hand-on lab exercises.

Student Take away

- Study Material
- Learning stuff
- Sample project for practice

Hadoop Administration online training curriculum

Understanding Big Data Ecosystem

- Introduction to Big Data
- Need for Big Data ecosystem
- Batch + Streaming

Understanding Hadoop

- Understanding the need for Hadoop
- Introduction to Hadoop ecosystem

Hadoop Architecture and Daemons

- Hadoop Architecture
- Various support Daemons

YARN and its daemons

- Resource Manager
- Node Manager
- RM queues
- Capacity and Fair scheduler

Map Reduce

- Understanding Sample MR code
- MR tuning parameters

HDFS and its features

- Name Node architecture
- Data node and Heartbeats

Cluster planning and Management

- Hadoop Tuning and planning

HDFS HA

- Journal based HA
- NFS based HA

YARN HA

Apache Ambari and Cloudera Manager

Apache Ambari

- Automated installation using Apache Ambari
- Ambari stacks
- Ambari Blueprints

Cloudera Manager

- Cloudera templates
- Hadoop and Ecosystem components

Hadoop Security

- Apache Ranger
- Kerberos
- Apache KNOX