

BIG DATABASE – HADOOP



Overview

Hadoop Training course teaches experienced / knowledge peoples on purpose of Hadoop Technology, how to setup Hadoop Cluster, how to store BigData using Hadoop (HDFS) and how to process/analyze the BigData using Map-Reduce Programming or by using other Hadoop ecosystems.

Prerequisite

- Basic Unix Commands
- Core Java (OOPS Concepts, Collections , Exceptions) — For Map-Reduce Programming
- SQL Query knowledge – For Hive Queries

Hardware and Software Requirements

- Any Linux flavor OS (Ex: Ubuntu/Cent OS/Fedora/RedHat Linux) with 4 GB RAM (minimum), 100 GB HDD
- Java 1.6+
- Open-SSH server & client
- MYSQL Database
- Eclipse IDE
- VM Ware (To use Linux OS along with Windows OS)

Course Duration

50 Hours, daily 1:30 Hours

Course Contents

1. Introduction to Hadoop
2. Hadoop Framework Architecture
3. Starting with Hadoop and Hadoop Program Building
4. Pig
5. Pig Advanced
6. Hive
7. HBase
8. Sqoop
9. ZooKeeper
10. Hadoop Administration
11. Built up over Hadoop
12. Hadoop in-Depth

Course Contents Details

1. Introduction to Hadoop

- What is Hadoop
- Why we need Hadoop
- Understand Big Data
- HDFS And Map Reduce

2. Hadoop Framework Architecture

- Block
- Rack
- Name Node
- Data Node
- Job Tracker
- Task Tracker
- Map Reduce
- Hadoop File system (HDFS)

3. Starting with Hadoop and Hadoop Program Building

- Apache Hadoop
- Hadoop Requirement
- Get Hadoop in your PC
- Design Control Flow
- Mapper
- Reducer

4. Pig

- What is Pig
- Pig Installation And Configuration
- Writing Simple Pig Programs

5. Pig Advanced

- Writing Complex Pig Programs
- User Defined Functions in Pig
- Optimization in Pig

6. Hive

- What is Hive
- Hive Installation And Configuration
- Writing Simple to Complex Hive Queries

7. HBase

- What is HBase

- HBase Installation And Configuration
- Understand the Basics of HBase
- HBase Vs RDBMS
- HBase– A Practical Approach

8. Sqoop

- What is Sqoop
- Sqoop Installation And Configuration
- Understand the Basics of Sqoop
- Import And Export in Sqoop
- Sqoop– A Practical Approach

9. ZooKeeper

- What is ZooKeeper
- ZooKeeper Installation And Configuration
- Understand the Basics of ZooKeeper
- Administration through ZooKeeper

10. Hadoop Administration

- WEB UI
- Administer Map Reduce
- Administer Name Node
- Administer Data Node
- Administer Job Tracker
- Administer Task Tracker
- Remove Node
- Assign Node
- Scheduling and Debugging Job

11. Built up over Hadoop

- Ambari
- Avro
- Spark
- Chukwa
- Mahout
- Cassandra
- Data Warehousing
- Mining

12. Hadoop in-Depth

- A Very Close Lookout
- Performance Tuning in Hadoop