

# RedBush Technologies

209 ,Unitech Arcadia ,South City 2 Sector 49 , Gurgaon.122018

PH : 9971483222 / 9718212607

## HadoopTraining

### Hadoop Training Course Duration

- Approx. 40 Hours

### Hadoop Course Content

#### Big Data and Hadoop

- Introduction to big data and Hadoop
- Hadoop Architecture
- Hadoop Versioning and Configuration
- Linux commands and Hadoop commands
- Cluster architecture and block placement
- Modes in Hadoop
- Hadoop Daemon
- Task Instance
- Hadoop Distributed File System
- HDFS Commands

#### Map-Reduce

- Understanding Map Reduce Framework
- Inspiration to Word-Count Example
- Developing Map-Reduce Program using Eclipse Luna
- HDFS Read-Write Process
- Map-Reduce Life Cycle Method
- Serialization(Java)
- Datatypes
- Comparator and Comparable(Java)
- Custom Output File
- Analysing Temperature dataset using Map-Reduce
- Custom Partitioner & Combiner
- Running Map-Reduce in Local and Pseudo Distributed Mode.

#### Hadoop Eco Systems

##### Hive

- Hive Introduction & Installation
- Data Types in Hive
- Commands in Hive
- Exploring Internal and External Table
- Partitions
- Complex data types
- UDF in Hive
- Thrift Server
- Java to Hive Connection

- Joins in Hive
- Working with HWI
- Bucket Map-side Join
- More commands
  1. View
  2. SortBy
  3. Distribute By
  4. Lateral View
  5. Running Hive in Cloudera

## **PIG**

- Installation and Introduction
- WordCount in Pig
- NYSE in Pig
- Working With Complex Datatypes
- Pig Schema
- Miscellaneous Commands
  1. Group
  2. Filter
  3. Order
  4. Distinct
  5. Join
  6. Flatten
  7. Co-group
  8. Union
  9. Illustrate
  10. Explain
- UDFs in Pig
- Parameter Substitution and DryRun
- Pig Macros

## **HBase**

- Import and Export Wizard
- HBase Introduction & Installation
- Exploring HBase Shell
- HBase Storage Technique
- HBasing with Java
- CRUD with HBase
- Hive HBase Integration