Lecture will be taken by an expert in Hadoop with overall 7 years of experience in IT industry

Demo lecture will be conducted initially of 2 hours. (**10 a.m-12 p.m**)

Topic -: **Hadoop/MapReduce/HDP/Cloudera/Microsoft Azure** (Cloud Computing)

Duration -: **4hrs per Day (32-40 hrs.)**

Days -: Saturday and Sunday (2 days a week)

Time -: 10:00 am to 02:00 pm (**Negotiable**)

Course Ware -: Theory + Hands on Practical

Course Syllabus will be as below -:

1. **Big Data Overview**
   - 1.1 Big Data Definition
   - 1.2 What is Big Data & What Comes Under Big Data
   - 1.3 Benefits of Big Data
   - 1.4 Big Data Technologies
2. **Hadoop Overview**
   - 2.1 History of Hadoop
   - 2.2 Hadoop Architecture
   - 2.3 Introduction to Hadoop Framework
   - 2.4 HDFS – Hadoop Distributed File System
   - 2.5 fIntroduction to MapReduce Framework
   - 2.6 Hadoop Work Flow
   - 2.7 Advantages of Hadoop
3. **Hadoop Environment & Cluster Setup**
   - 3.1 Hadoop Cluster Administrator Roles & Responsibilities
   - 3.2 Pre-requisite for Hadoop
   - 3.3 Installation of Java
   - 3.4 Download & Install Hadoop
   - 3.5 Hadoop Operation Modes
   - 3.6 Deploy Hadoop in Standalone Mode

3.7 Deploy Hadoop in Pseudo Distributed Mode
3.8 Verifying Hadoop Installation
3.9 Understanding of Namenode, Datanode, Job Tracker & Task Tracker

4. HDFS Overview & Operations

4.1 Features & Goals of HDFS
4.2 HDFS start up
4.3 List, Insert, Retrieve Data from HDFS
4.4 Shutting Down HDFS
4.5 HDFS CLI

5. MapReduce Overview & Operations

5.1 What is MapReduce
5.2 MapReduce Algorithm & Stages
5.3 Running MapReduce Jobs
5.4 Terminologies

6. Hadoop Cluster Planning & Managing

6.1 Planning Hadoop Cluster
6.2 Hardware/Software Considerations
6.3 Rack Awareness

7. HDP Setup

7.1 Introduction to HDP
7.2 Hadoop Cluster Installation Using Ambari
7.3 Cluster Monitoring/Troubleshooting
7.4 Types of Schedulers in Hadoop – FIFO, Capacity & Fair
7.5 Setup Queues & Pools for Jobs
7.6 Managing/Scheduling Jobs
7.7 Look at Performance Tuning Parameters

8. Cloudera Setup

8.1 Introduction to Cloudera
8.2 Hadoop Cluster Installation Using Cloudera manager
8.3 Cluster Monitoring/Troubleshooting

9. Backup, Recovery & Maintenance

9.1 Add Storage to Datanode
9.2 Setting up Stand byNamenode
9.3 Hadoop Backup
9.4 Whitelist & Blacklist Datanodes
9.5 Hadoop Balancer
9.6 Diagnostics & Recovery

10. YARN Overview & Operations

10.1 YARN Configuration
10.2 YARN Execution
10.3 YARN Workflow

11. Big Data Components Overview

11.1 Introduction & Installation of HIVE
11.2 Introduction & Installation of HBase

11.3 Introduction & Installation of Sqoop
11.4 Introduction to Oozie

12. Microsoft Azure

12.1 Introduction to Cloud Computing
12.2 Introduction to Azure
12.3 Virtual Machine creation on Azure
12.4 Virtual Machine management on Azure

A YEAR FROM NOW YOU WILL WISH YOU HAD STARTED TODAY…!!!

Hope to see you again…!!!

Contact Details

Rajiv S Banerji

Mumbai

Q-13, Row House Type 2 (RH-2), SAROJ NIWAS

Sector 6, Vashi, New Bombay 400703.

Land Mark : Near Dena Bank

Bus Stop :VashiGaon Stop

(Opp Wockhardt Hospital on Highway) Just after Toll Naka

Railway : 20 mins walking from Vashi station

Mobile : 9820851665

Email : rajiv@opensourceforce.com

Note -:

1. Practical will be performed on Azure Cloud (Free).
2. Please bring laptop for practice.
3. Students who do not have laptops, provision will be made. Later on they can do the practical on their desktop at home.