Big Data Developer Training Program

(40 Hours of training – Prepare of HDPCA Certification – Hadoop Administration)

Day 1:

Introductions

    a) Meet with your Instructor
    b) Course Overview
    c) Administration Roadmap
    d) Prerequisites

Lesson 1 – Introduction to Big Data

    a) Why BigData ? - why the rush?
    b) Tradition architecture Vs Hadoop Architecture
    c) Elements of Hadoop Ecosystem
    d) Pros and Cons
    e) Hadoop Distribution

Lesson 2 - Introduction to Linux / CentOS

    a) Basic Linux commands (ls, yum, wget, vi, view, cat, cp, rm, df –h, sudo su -, useradd, usermod etc)
    b) LAB 1
    c) Cluster pre-requisite
    d) Node preparation
    e) Deciding on master and slave node

Lesson 3 - Setting up cluster

    a) Firewall consideration, Ports and Public IP discussion
    b) Installing Ambari-Server
    c) Installing Ambari-agent
    d) Adding Master and Slave Node to the configuration
    e) Adding Master and Slave daemons
    f) Setting up Bastion / EdgeNode / Gateway node
    g) LAB and Excercises

Lesson 4 – Component installation and configuration

    a) HDFS, YARN, MapReduce
    b) Hive, Pig, Tez
    c) Sqoop
    d) Flume, Storm
    e) Ranger
    f) MySql
    g) Oozie, Zookeeper

Lesson 5 – Day to Day Operation

Big Data Developer Training Program

User space creation and management

a. Unlike Linux there are No user in HDFS, rather its user space
b. If there are matching user directory in Linux with a directory in HDFS, we say that's it's a user space for that user
c. Managing space between application users
d. Best Practice – Create a user account and matching user-space for department
e. Setup user account in Ambari
f. Managing Ambari users, groups and permissions
g. Assign Space and name quotas
h. HDFS SnapShot
i. Demo
j. Distcp – Data backups and migration

Lesson 6 - Adding new hosts, services and components

Lesson 7 - Monitoring using Ambari – Alerts and Notification

Lesson 8 - Cluster Capacity management

Lesson 9 - Security – HDFS ACL's, Ranger, SSL/TLS

Lesson 10 - EdgeNode / Gateway Node configuration and Access

Lesson 11 – Lab Excercises

Lesson 19 - Exam Tips and Tricks