

SEVENMENTOR TRAINING PVT.LTD

Hadoop Administration Syllabus

ABOUT BIG DATA & HADOOP PROGRAM

Apache Hadoop is an open-source software framework used for distributed storage and processing of dataset of big data using the MapReduce programming model. It consists of computer clusters built from commodity hardware. All the modules in Hadoop are designed with a fundamental assumption that hardware failures are common occurrences and should be automatically handled by the framework. The core part of Hadoop consist of a Storage part, known as Hadoop Distributed File system(HDFS),and a processing part which is a mapreduce programming model.

Apache Hadoop clusters which includes support for Hadoop HDFS, Hadoop MapReduce, Hive, HBase, ZooKeeper, Oozie, Pig, and Sqoop.

Career Opportunities:

Companies like Google, EMC Corporation, Yahoo, Apple, HortonWorks, Oracle, Amazon, Cloudera, IBM, Cisco, Microsoft and many more have opened their doors for hadoop professionals. Various positions like product managers, hadoop developers, testers, database administrators, senior hadoop developers and alike are open. Companies are searching for experienced candidates as well as freshers.

Course Structure and Pre-Requisites

1. Software Developers/Programmers/Engineers, who are into Database/Programming and exploring for great job opportunities in Hadoop
2. Managers, who are looking for the latest technologies to be implemented in their organization, to meet the current and upcoming challenges of data management
3. Any Graduate/Post-Graduate, who is aspiring a great career towards the cutting edge technologies

Office 21-25/A, First Floor Shreenath Plaza,
Dnyaneshwar Paduka Chowk, Pune, Maharashtra 411005
Mob : +91-7798058777

SEVENMENTOR TRAINING PVT.LTD

Architecture of Hadoop and relates how HDFS carry data storage and YARN,the other core constituent of Hadoop, gives the data processing capacity.

Hadoop administration comprises how to form and configure a single node, multinode homogenous clusters.

Data replication is the important part of HDFS, and we will practice about how HDFS arrange its data and how data replication works.

As an Hadoop administrator,important job is to arrange HDFS storage.

1)Hadoop MultiNode Cluster Setup and Running a BigData Example

- Hadoop single node cluster
- Hadoop MultiNode Cluster

- Setup Three Node Hadoop cluster
- Running web Application on cluster
- Analyze the cluster using
 - a)NameNode UI
 - b)JobTracker UI

- SettingUp Replication Factor

Installation and configuration of tools

E)HIVE

F)PIG

G)SQOOP

H)Oozie

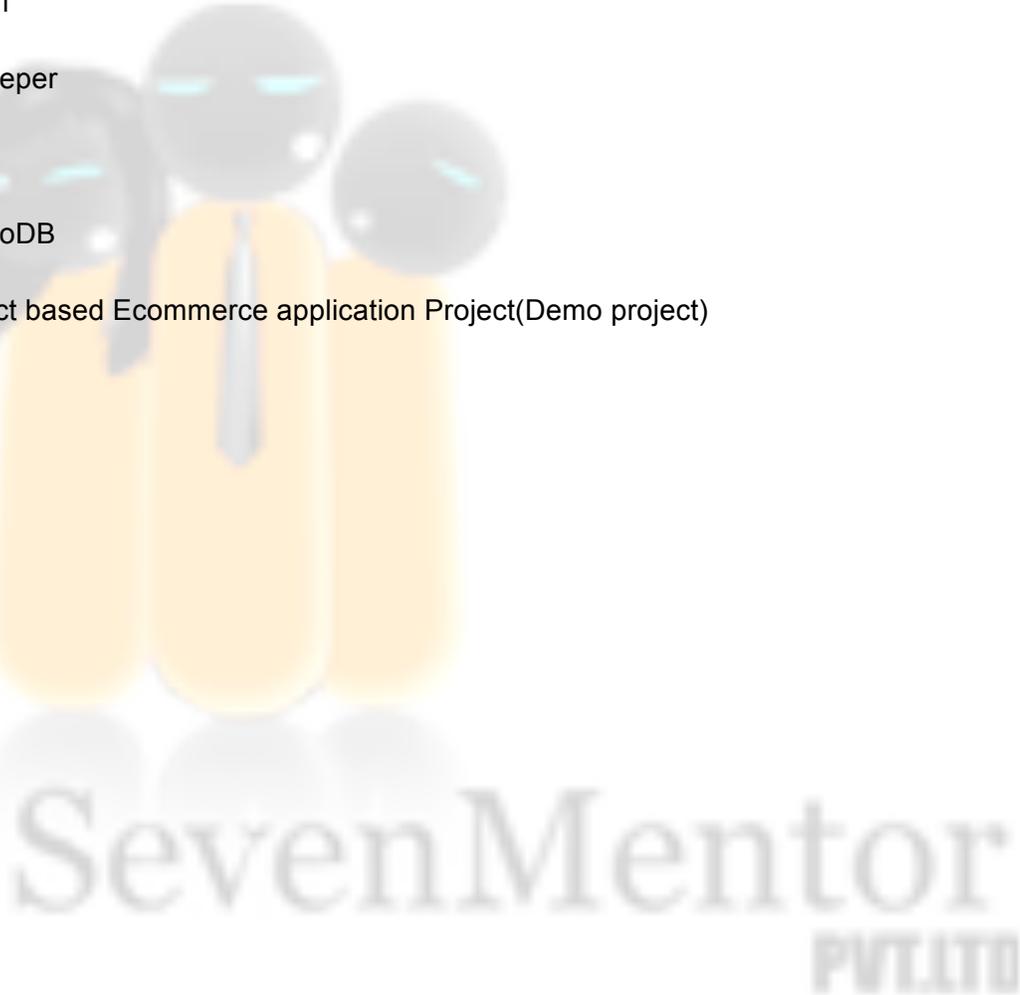
I)HBASE

J)Flume

Office 21-25/A, First Floor Shreenath Plaza,
Dnyaneshwar Paduka Chowk, Pune, Maharashtra 411005
Mob : +91-7798058777

SEVENMENTOR TRAINING PVT.LTD

- K)Mahout
- M)Solr
- N)Ambari
- O)Zookeeper
- P) Spark
- Q) MongoDB
- R)Product based Ecommerce application Project(Demo project)



SevenMentor
PVT.LTD

Office 21-25/A, First Floor Shreenath Plaza,
Dnyaneshwar Paduka Chowk, Pune, Maharashtra 411005
Mob : +91-7798058777