HADOOP SCHOOL OF TRAINING
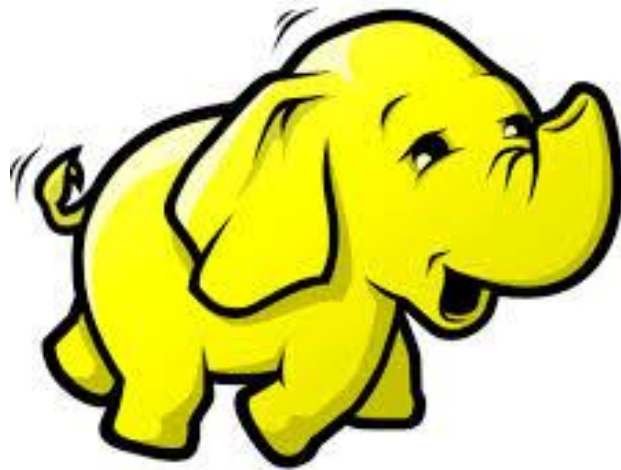

Hadoop Development & BI- 0 to 100

HADOOP SCHOOL OF TRAINING

Hadoop Development - 0 to 100

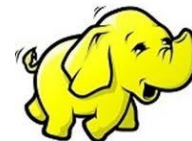| Basics	Hands On	Development
Learn the basics of Big Data and hadoop	Play with Hadoop and hadoop ecosystem	Become a top notch hadoop developer

Hadoop Development Program- 0 to 100 (40 Hours)

Overview of the course:

Hadoop Development Program is a one stop course that introduces you to the domain of hadoop development as well as gives you technical knowhow of the same. At the end of this course you will be able to earn a credential of hadoop developer and you will be capable of dealing with Terabyte scale of data and analyze it successfully using mapreduce
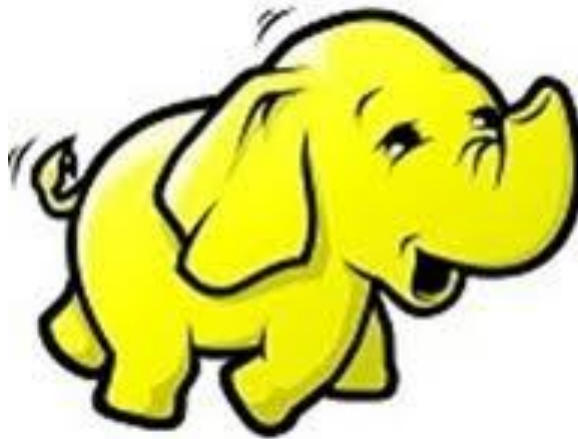
Who this course is for and not for?

For: Typically professionals with basic knowledge of software development, programming languages, and databases will find this course really helpful. Basic knowledge should be enough to succeed at this course

 Not For: Students who are absolute beginners at software development as a discipline will find it difficult to follow the course

www.hadoopschooloftraining.com; Ph: 9325793756; Email:learninghub01@gmail.com
2nd Floor, Above HDFC Bank, Next to Noble Polyclinic, MAGARPATTA CITY, PUNE

HADOOP SCHOOL OF TRAINING

Phase 1: Hadoop Fundamentals (20 Hours)

Getting the Basics Rights

Big Data

- What is Big Data
- Dimensions of Big Data
- Big Data in Advertising
- Big Data in Banking
- Big Data in Telecom
- Big Data in eCommerce
- Big Data in Healthcare
- Big Data in Defence
- Processing options of Big Data
- Hadoop as an option

Hadoop

- What is Hadoop
- How Hadoop Works
- HDFS

- Mapreduce
- How Hadoop has an edge

Hadoop Ecosystem

- Sqoop
- Oozie
- Pig
- Hive
- Flume

Hadoop Hands On

- Setting up Hadoop on a Single node cluster
- Running HDFS commands
- Running your Mapreduce program
- Running Sqoop Import and Sqoop Export
- Creating Hive tables

Directly from Sqoop
- Creating Hive tables
- Querying Hive tables
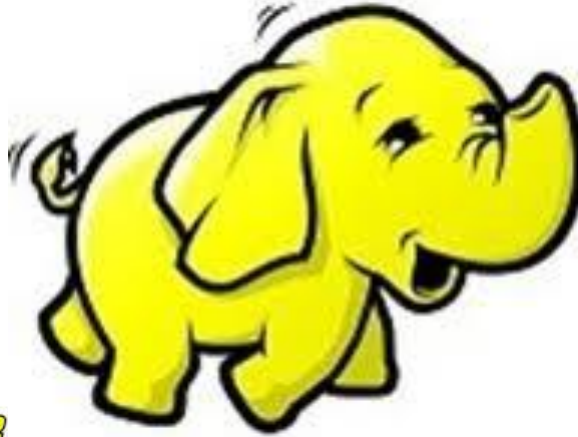- Running an Oozie workflow
- Analyzing twitter data using Flume

Multinode Setup

- Setting up Multinode setup on Amazon ec2
- Setting up multimode setup on the classroom machines
- Setting up Cloudera Manager on the cloud
- Setting up Cloudera Manager on local setup

Cluster Capacity Planning

Level 1: Mini Project

Level 1: Evaluation Test (50 marks)

HADOOP SCHOOL OF TRAINING

Phase 2: Hadoop Development
(12 hours)

Become a Pro developer

Advanced Mapreduce

- Mapreduce Code Walkthrough
- ToolRunner
- MR Unit
- Distributed Cache
- Combiner
- Partitioner
- Setup and Cleanup methods
- Using Java API to access HDFS
- Map Side joins
- Reduce side joins
- Input Types in Mapreduce
- Output Types in Mapreduce
- Custom Input Data types
- Custom Output Data types
- Multiple reducer MR program
Zero Reducer Mapper Program

Advanced Mapreduce Hands On

- MR Unit hands On
- Distributed Cache hands On
- Partitioner hands On
- Combiner hands On
- Accessing files using HDFS API hands on
- Map Side joins hands on
- Reduce side joins hands on

Mapreduce Design Patterns:

- Searching
- Sorting
- Filtering
- Inverted Index
- F-IDF
- Word Co-occurrence

Mapreduce Design Patterns Hands On:

- Searching Hands On
- Sorting Hands On
- Filtering Hands On
- Inverted Index Hands On
- TF-IDF – Hands On
- Word Co-occurrence

HADOOP SCHOOL OF TRAINING

Phase 3: Hadoop BI
(8 hours)

Analyze data using Pig and Hive

Pig
- Introduction
- Basic Data Analysis
- Pig Hands on

Hive
- Introduction
- Basic Data Analysis with Hive
- Hive Data Management
- Hive Hands On

Evaluation Test (50 marks)

HADOOP SCHOOL OF TRAINING

Trainer Profile

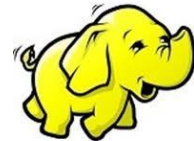| Experienced	Certified	References Customer
8+ yrs of Enterprise Software Dev Exp.	Hadoop, Hbase and MapR certified	Accenture, HP, Genpact, Mastek, and Cisco, Mphasis, Deluxe Etertainment

About the trainer

Trainer's Certifications

| CCAH,CCHD, CCHSB | MapR M5 | Zend | SCJP | SCWCD |