Become a Cloudera Certified Developer for Apache Hadoop

Why Big Data and Hadoop 2.0

The use of Big Data is becoming a crucial way for leading companies to outperform their peers. It enables us to tap into large information flow where data about products and services, buyers and suppliers, consumer preferences and intents can be captured and analyzed. Therefore nurturing Big Data Analytics expertise plays an essential role in developing new products or services.

Who should attend this Hadoop 2.0 training

This training is ideal for professionals who wish to implement Big Data Analytics by using Hadoop framework.

What you will learn on Hadoop 2.0

The course equips participants to work on the Hadoop environment with ease and learn vital components such as Flume, Hive, Pig, Map Reduce and other advanced concepts like Hadoop 2.0: HBase, ZooKeeper and Sqoop.

The course consists of:
3 Days Classroom training
20 Hours of Real Time Industry based Projects
15 Hours of Lab Exercises with proprietary VM
3 Hours of Doubt Clarification Session
2 Projects with Unique Data Sets Included
Industry Specific Projects on the Retail and Telecom Sectors
1 Big Data & Hadoop Simulation Exam
Downloadable e-Book Included
Java Essentials for Hadoop Included
Hadoop Installation Procedure Included
Excellence in Hadoop Certificate
Hadoop Deployment and Maintenance Tips
Packed with Latest & Advanced modules like Yarn, Flume, Oozie, Mahout, and Chukwa
Tips and Techniques to clear Certification Exam by the Trainer

How you will be certified

Certificate 1: Trainer will provide guidance by going through the practice test and answer with explanations. Participant will sit for web-based CCDH (Cloudera Certified Developer for Apache Hadoop) examination and receive a certificate upon scoring 80% or above in the examination.

Certificate 2: Bring back what you have learnt in the 3-day course and implement the Hadoop ecosystem into your product or services by completing a Hadoop project. Trainer will assist and provide guidance via web-based correspondence. Participant will receive another certificate upon completion of project.

Date: 27-29 Jul 2015
        (Monday - Wednesday)
Time: 0900 - 1700
Venue: Suite 2B-21-1, Level 21, Block 2B,
        Plaza Sentral, Jalan Stesen Sentral 5,
        KL Sentral, 50470 Kuala Lumpur,
        Malaysia.

HRDF Claimable

Course Outline

Day 1

Introduction to Big Data
➢ The driving force of the Big Data era
➢ Why Big Data deserves your attention
➢ Use cases of Big Data
➢ Different options of analyzing Big Data
➢ Introduction to Hadoop
➢ History of Hadoop
➢ Problems with traditional large-scale systems and need for Hadoop
➢ Where Hadoop is being used
➢ Understanding distributed systems

RDBMS and Hadoop

Understanding Hadoop Architecture
➢ Understanding Hadoop configuration files
➢ Hadoop components- HDFS, MapReduce
➢ Overview of Hadoop processes
➢ Overview of Hadoop Distributed File System
➢ NameNode
➢ DataNode
➢ The Command-Line Interface

Hands-On Exercise
➢ Using HDFS commands

Day 2

Introduction to Apache Pig
➢ Pig architecture
➢ Data types, Arithmetic, Relational Operators, Pig UDF
➢ Pig built in functions: Eval, Math, String, Date, Time functions
➢ Pig assignment

Introduction to Apache Hive
➢ Hive architecture
➢ Hive Metastore configuration
➢ Data types of file formats and Hive built in/basic functions
➢ Databases in Hive
➢ External and internal tables, and loading data into Hive

Hands-on Exercise
➢ Pig and Hive

Day 3

Hive
➢ Bucketing and partitions
➢ HiveQL queries: Select, Where, Group By, Having, Distributed By with Sort By, Cluster By, Order By vs Sort By
➢ Hive UDF

MapReduce Program In Java using Eclipse

No-SQL Introduction
➢ Hive
➢ Differences between SQL and No-SQL
➢ HBase architecture
➢ ZooKeeper introduction

Hands-on Exercise
➢ HBase
➢ Sqoop and Flume,HBase
➢ Hadoop Admin - Hadoop Installation

About Go Training

Go Training applies effective pedagogical methodologies that demonstrate case studies and hands-on practical skills, in addition to explaining clearly how things work in principle. Every course that we conduct is delivered by a subject matter expert who holds the academic qualification and working experience in that specialization. On the days when they are not teaching, our trainers work on consultancy projects and technical deliveries. Their work has received numerous recognition and awards in the industry. Our team of trainers has been invited as keynote speakers at numerous international conferences, and as principal consultants for various industries.

About the Instructor

Mr. Dhana Naidu graduated from Perivar University with Master of Computer Application degree. He previously worked for Emerio (an NTT Communications) Corporation Pvt Ltd as a Hadoop Developer and for Verilog Networks Pvt Ltd as software engineer. He is currently the Hadoop Developer with IBM in Bangalore, India. He is a Cloudera Certified Developer for Apache Hadoop (CCDH) and cloudera certified administrator for apache hadoop (CCAH). He has three years of industrial and hands-on experience working with HDFS, MapReduce, Pig, Hive, HBase, Sqoop, Flume and Oozie, as well as more than two years of exposure in Data Base Test Analyst (Microsoft Sql Server 2008). He was awarded the "Nokia Siemens Networks Best Contributor" at Emerio Corporation (an NTT Communications) in 2013 which was an appreciation from the client, Nokia Siemens Networks for significant contribution in the project which involved storing millions of mobile call records and providing real time access to the call records and billing information to customers, and analysis of such large volumes of data leads using Big Data Concepts and Hadoop Framework.