



Caliber Data Training
In partnership with
MetaScale



MetaScale Hadoop Developer Training

Course Description

This hands-on course gives students an introduction to Big Data and covers Apache Hadoop core concepts including the Map/Reduce paradigm. It features real-world use cases and examples to illustrate this powerful framework. The course has been designed by MetaScale's practicing professionals who have deployed Hadoop technologies at data-intensive enterprises.

Who Should Attend

This course is intended for developers who want to learn more about Hadoop technologies. Beginners will learn extensively about one of the hottest technologies in the IT market today. Students who are familiar with Hadoop will improve their skills by learning more advanced features of the product.

Duration

4 days

Prerequisites

The students should have:

- Basic Linux and Shell scripting
- Knowledge of RDMS and basic SQL
- Programming knowledge, preferably in Java.

Course Topics

- Apache Hadoop and ecosystem Overview
- HDFS
- Map Reduce
- PIG
- Hive
- Sqoop

Course Outline

- I. APACHE HADOOP AND ECOSYSTEM OVERVIEW**
 - A. Introduction to Big Data
 - B. The Motivation for Hadoop
 - C. Brief History of Hadoop
 - D. Hadoop Core Concepts
 1. HDFS
 2. M/R
 3. Name Node
 4. Data Node
 5. Job Tracker
 6. Task Tracker



Caliber Data Training
In partnership with
MetaScale



-
- E. Hadoop Ecosystem
 - 1. Sqoop
 - 2. Pig
 - 3. Hive
 - 4. Hue
 - 5. Brief Introduction to Other Components

II. HDFS

- A. Introduction to HDFS
- B. HDFS Architecture Overview
- C. HDFS Concepts
- D. HDFS Commands

III. MAP REDUCE

- A. M/R Paradigm
- B. M/R Phases Overview and Architecture
- C. Use Cases/ Examples
- D. M/R Code

IV. PIG

- A. Introduction
- B. Pig Architecture
- C. Pig Latin Examples
- D. Commands to Manipulate Data
- E. Built in Functions
- F. Examples

V. HIVE

VI. SQOOP

m/e13