

Course Objective Summary

- ✓ During this course, you will learn:
- ✓ Introduction to Big Data and Hadoop
- ✓ Hadoop ecosystem - Concepts
- ✓ Hadoop Map-reduce concepts and features
- ✓ Developing the map-reduce Applications
- ✓ Pig concepts
- ✓ Hive concepts
- ✓ Oozie workflow concepts
- ✓ HBASE Concepts
- ✓ Real Life Use Cases
- ✓ HDFS
- ✓ Map Reduce
- ✓ Statistics

Introduction to Big Data and Hadoop

- ✓ What is Big Data?
- ✓ What are the challenges for processing big data?
- ✓ What technologies support big data?
- ✓ What is Hadoop?
- ✓ Why Hadoop?
- ✓ History of Hadoop
- ✓ Use Cases of Hadoop
- ✓ Hadoop eco System

Understanding the Cluster

- ✓ Typical workflow
- ✓ Writing files to HDFS
- ✓ Reading files from HDFS
- ✓ Rack Awareness
- ✓ 5 daemons

Let's talk Map Reduce

- ✓ Before Map reduce
- ✓ Map Reduce Overview
- ✓ Word Count Problem
- ✓ Word Count Flow and Solution
- ✓ Map Reduce Flow
- ✓ Algorithms for simple & Complex problems

Developing the Map Reduce Application

- ✓ Data Types
- ✓ File Formats

Mind Q Systems Private Limited

- ✓ Explain the Driver, Mapper and Reducer code
- ✓ Configuring development environment - Eclipse
- ✓ Writing Unit Test
- ✓ Running locally
- ✓ Running on Cluster
- ✓ Hands on exercises

How Map-Reduce Works

- ✓ Anatomy of Map Reduce Job run
- ✓ Job Submission
- ✓ Job Initialization
- ✓ Task Assignment
- ✓ Job Completion
- ✓ Job Scheduling
- ✓ Job Failures
- ✓ Shuffle and sort
- ✓ Oozie Workflows
- ✓ Hands on Exercises

Map Reduce Types and Formats

- ✓ MapReduce Types
- ✓ Input Formats - Input splits & records, text input, binary input, multiple inputs & database input

- ✓ Output Formats - text Output, binary output, multiple outputs, lazy output and database output
- ✓ Hands on Exercises

Map Reduce Features

- ✓ Counters
- ✓ Sorting
- ✓ Joins - Map Side and Reduce Side
- ✓ Side Data Distribution
- ✓ MapReduce Combiner
- ✓ MapReduce Partitioner
- ✓ MapReduce Distributed Cache
- ✓ Hands Exercises

Hive and PIG

- ✓ Fundamentals
- ✓ When to Use PIG and HIVE
- ✓ Concepts
- ✓ Hands on Exercises

HBASE

- ✓ CAP Theorem
- ✓ Introduction to NOSQL
- ✓ Hbase Architecture and concepts
- ✓ Programming and Hands on Exercises

Mind Q Systems Private Limited



Case Studies Discussions

Certification Guidance

Mind Q Systems Pvt Ltd

Ph: +91.40.66 66 42 91 / 92

Email: info@mindqsystems.com; Url: www.mindqsystems.com; : www.facebook.com/mindqsystems