

Trusted Mentor in Your True Success

Join us Today, Get Trained by Industry Experts with Live Industry Exposure

Course Description

Administrator Training course for Apache Hadoop provides participants with a comprehensive understanding of all the steps necessary to operate and maintain a Hadoop cluster. The course topics include Introduction to Hadoop and its Architecture, MapReduce and HDFS and MapReduce Abstraction. From installation and configuration through load balancing and tuning, this training course is the best preparation for the real-world challenges faced by Hadoop administrators. It further covers best practices to configure, deploy, administer, maintain, monitor and troubleshoot a Hadoop Cluster.

After completing this course, student will be able to:

- Understand Hadoop main components and Architecture
- Deep dive into Hadoop Distributed File System (HDFS)
- Understand concepts of Yarn
- Understand MapReduce abstraction and its working
- Plan and Deploy a Hadoop cluster
- Optimize Hadoop cluster for high performance, based on specific job requirements
- Monitor Hadoop cluster and Execute routine Administration procedures
- Handle Hadoop component failures and recoveries
- Determining the correct hardware and infrastructure for your cluster
- How to load data into the cluster from dynamically-generated files using Flume and from RDBMS using Sqoop
- Configuring the Fair Scheduler to provide service-level agreements for multiple users of a cluster
- Best practices for preparing and maintaining Apache Hadoop in production
- Troubleshooting, diagnosing, tuning, and solving Hadoop issues

Who Should Attend

This course is best suited to systems administrators and IT managers who have basic Linux experience. Fundamental knowledge of any programming language and Linux environment. Participants should know how to navigate and modify files within a Linux environment. Prior knowledge of Apache Hadoop is not required.

Modules Covered

- Introduction
- Hadoop Cluster Installation
- The Hadoop Distributed File System (HDFS)
- MapReduce and Spark on YARN
- Hadoop Configuration and Daemon Logs
- Getting Data Into HDFS
- Planning Your Hadoop Cluster
- Installing and Configuring Hive, Impala, and Pig
- Hadoop Clients Including Hue
- Advanced Cluster Configuration
- Hadoop Security
- Managing Resources
- Cluster Maintenance
- Cluster Monitoring and Troubleshooting

Course Duration

- **Regular Track** : 5 Weeks (2 Hours/Day)
- **Weekend Track** : 5 Weekends (4 Hours/Day)

Kits Include

- Pre-course technical evaluation
- SI-Android Tools Kit
- Training Material (E-Books)
- Certificate of Excellence from CODEC Networks

Post Training Program (CODEC Networks Specialty)

- Live Project Work
- Hand-over Labs & Practical's Checklist for review
- Placement Assistance **
- Discount Vouchers up to 15 - 25% for further training