



# Hadoop Admin Training Curriculum

## STRUCTURE



## Big Data Hadoop Admin Training Curriculum

*“Hadoop Admin Certification Training to make your career more rewarding and demanding.”*

### **Course Objectives:**

- Get yourself signed in for our comprehensive; real-world projects led Hadoop Training Program to master the skills and tools in the Hadoop ecosystem.
- Gain all the practical learnings around Hadoop Architecture, Apache PIG, Hive, HBase, YARN that will help you in qualifying the competent Hadoop Certifications and make your career a demanding one.
- Prepare yourself for the Hadoop admin global certification exam and get recognition
- Apply for leading MNCs and get hired by top industries across the world.

### **Course Description:**

Croma Campus offers the best Hadoop admin Training in Noida with most experienced professionals. Our Instructors are working in Big Data Space and related technologies for years in MNC's.

We aware of industry needs and we are offering Hadoop admin Training in more practical way. Our team of Hadoop trainers offers the in-classroom training with best industry practices.

We framed our syllabus to match with the real-world requirements for both beginner level to advanced level. Our training will be handled in either weekday or weekends programme depends on participants requirements.

Further, the course content is prepared keeping latest industry trends and certification structure in mind. To know more, you can contact our expert team and get more details.

### **Course Content:**

#### **Module 1: Introduction to Big Data & Hadoop**

- Introduction to Big Data
  - Overview of Course
  - What is Big Data?
  - Big Data Analytics
  - Challenges of Traditional System
  - Distributed Systems
- Introduction to Hadoop
  - Components of Hadoop Ecosystem
  - Commercial Hadoop Distributions
  - Why Hadoop?
  - Fundamental Concepts in Hadoop

## **Module 2: Hadoop Installation and Initial Configuration**

- Deployment Types
- Installing Hadoop
- Specifying the Hadoop Configuration
- Performing Initial HDFS Configuration
- Performing Initial YARN and MapReduce Configuration
- Hadoop Logging

## **Module 3: Hadoop Security**

- Why Hadoop Security Is Important
- Hadoop's Security System Concepts
- What Kerberos Is and How it Works
- Securing a Hadoop Cluster with Kerberos

## **Module 4: HDFS**

- HDFS Features
- Writing and Reading Files
- NameNode Memory Considerations
- Overview of HDFS Security
- Using the NameNode Web UI
- Using the Hadoop File Shell

## **Module 5: Installing and Configuring Hive, Impala and Pig**

- Hive
- Impala
- Pig

## **Module 6: Managing and Scheduling Jobs**

- Managing Running Jobs
- Scheduling Hadoop Jobs
- Configuring the Fair Scheduler
- Impala Query Scheduling

## **Module 7: Getting Data into HDFS**

- Ingesting Data from External Sources with Flume
- Ingesting Data from Relational Databases with Sqoop
- REST Interfaces
- Best Practices for Importing Data

## **Module 8: Hadoop Clients**

- What is a Hadoop Client?
- Installing and Configuring Hadoop Clients
- Installing and Configuring Hue
- Hue Authentication and Authorization

## **Module 9: Cluster Maintenance**

- Checking HDFS Status
- Copying Data between Clusters
- Adding and Removing Cluster Nodes
- Rebalancing the Cluster
- Cluster Upgrading

## **Module 10: YARN and MapReduce**

- What Is MapReduce?
- Basic MapReduce Concepts
- YARN Cluster Architecture
- Resource Allocation
- Failure Recovery
- Using the YARN Web UI
- MapReduce Version 1

## **Module 11: Cloudera Manager**

- The Motivation for Cloudera Manager
- Cloudera Manager Features
- Express and Enterprise Versions
- Cloudera Manager Topology
- Installing Cloudera Manager
- Installing Hadoop Using Cloudera Manager
- Performing Basic Administration Tasks using Cloudera Manager

## **Module 12: Cluster Monitoring and Troubleshooting**

- General System Monitoring
- Monitoring Hadoop Clusters
- Common Troubleshooting Hadoop Clusters
- Common Misconfigurations

## **Module 13: Planning Your Hadoop Cluster**

- General Planning Considerations
- Choosing the Right Hardware
- Network Considerations

- Configuring Nodes
- Planning for Cluster Management

## **Module 14: Advanced Cluster Configuration**

- Advanced Configuration Parameters
- Configuring Hadoop Ports
- Explicitly Including and Excluding Hosts
- Configuring HDFS for Rack Awareness
- Configuring HDFS High Availability

## **Module 15: Placement Guide**

- Tips to clear an Interview
- Common Interview questions and answers
- Hadoop Interview Questions and Answers
- Resume Building Guide
- Hadoop career roadmap and certifications
- Attempt for Hadoop Global Certification Exam
- Start applying for Jobs