



Who we are

Techeduxon has been a platform for global online technology education since 2015. We are now taking it up a notch higher by introducing ways of advanced learning to up-skill and cross-skill your profession with cutting-edge, customized programs. With our Top Tier IT & Enterprise training Courses, we enable you to the forefront & become 'Industry Ready' in this Advancing & Unforeseen Digital World with your up skill Innovations.

Course overview

The Hadoop admin training enables you to work with the versatile frameworks of the Apache Hadoop ecosystem. This course covers Hadoop installation and configuration, computational frameworks for processing Hadoop administrator activities, cluster management with Sqoop, Flume, Pig, Hive, Impala, and Cloudera.

- Introduction to Hadoop
- Hadoop Distributed File system
- Hadoop cluster setup
- Hadoop configurations and Hadoop Daemon Logs
- Hadoop Computational framework
- Hadoop Administration and Cluster maintenance
- Hadoop cluster planning scheduling through schedulers
- Hadoop clients and HUE interface
- Data ingestion in Hadoop cluster
- Hadoop Ecosystem services
- Hadoop security
- Hadoop cluster monitoring

Prerequisites

No prerequisites are required to take up this course. But having Basic knowledge of mathematics and statistics is beneficial.

Target audience

Systems Administrators, IT managers, IT administrators and operators, IT systems engineers, data engineers and database administrators, data analytics



administrators, cloud systems administrators, web engineers, and individuals who intend to design, deploy, and maintain Hadoop clusters.

Course Modules

Introduction to Hadoop

- Introduction to Hadoop
- Hadoop & its Ecosystem
- Hadoop Core components
- Hadoop & its capabilities

Hadoop Distributed File system

- Hadoop Distributed file System, its architecture, working & internals
- Learn about possibilities without HDFS
- Identify requirements to setup a Hadoop cluster

Hadoop Cluster Setup

- Need of Cluster Management Solution
- Choice of Installation methods like Automated/ Manual
- Linux machines setup like Virtualization & Cloud
- Hadoop Cluster Setup which includes Apache Hadoop V2 & Cloudera Distribution of Hadoop (CDH)
- Cloudera manager features and capabilities
- Working with Hadoop cluster, HDFS & data
- Working with management console/ UI (user interfaces) & Linux terminals.

Hadoop configurations and Hadoop Daemon Logs

- Configuration files, ports & properties that relate to functioning of Hadoop cluster.
- Explain how to manage Hadoop configuration
- How to deal with stale configurations
- Locate log files generated on hosts
- Filter information in log files



- Explain how to get diagnostic information from log files

Hadoop Computational framework

- Role of computational framework
- Explain YARN framework and concepts
- Describe YARN applications
- Describe YARN memory and CPU settings

Hadoop Administration and Cluster maintenance

- List the steps for managing services including adding, deleting, starting, stopping and checking status of services
- Explain the procedure to enable rack awareness
- List the steps to add, remove and move role instances and hosts
- Cite the challenges faced with the first version of Hadoop
- Explain the features in the second version that help overcome the challenges faced with the first version

Hadoop cluster planning scheduling through schedulers

- Scheduling concepts
- Identify schedulers & ways to manage resources using schedulers
- Queue management
- Describe FIFO, Fair Scheduler, and Capacity Scheduler

Hadoop clients and HUE interface

- How to install and configure Hadoop client
- Process of how HUE works
- How to install and configure HUE interface
- Process of how authentication and authorization is managed in HUE interface.

Data ingestion in Hadoop cluster



- Understand data ingestion
- Understand how flume and Sqoop works

Hadoop Ecosystem services

- Some of the services and open-source components that work within the Hadoop ecosystem
- How to configure Hive in different modes
- Explain the architecture of HBase and cite the advantages of using HBase
- Advantages of Hive
- Working with Apache Kafka
- Architecture of Apache spark

Hadoop security

- Identification of the different threat categories
- Security aspects for different nodes
- Operating system security
- Describe Kerberos and how it works
- Describe Service Level Authorization

Hadoop Cluster monitoring

- Cluster monitoring and ways to choose right monitoring solution
- Different categories of Hadoop metrics

Projects & Assignments

Our Expert trainers will provide the real time Projects & Assignments.



Help
Support

Trainings@techeduxon.com

www.techeduxon.com