

SM
Certificate in Hadoop
Big Data & Analytics
Duration: 6 Months

Introduction: Hadoop is an open-source framework that allows to store and process big data in a distributed environment across clusters of computers using simple programming models. It is designed to scale up from single servers to thousands of machines, each offering local computation and storage. The course is designed to give you in-depth knowledge of the Big Data framework using Hadoop including HDFS, YARN, and Map Reduce. You will learn to process and analyze large datasets stored in the HDFS.



Linux Operating System

Linux introduction | Linux file system | Partitioning the Hard drive for Linux | Installing the Linux system | Linux commands | Processes in Linux | Shell programming | awk programming | System administration | Common administrative task | KDE & Gnome graphical interfaces | Basic networking administration setting | TCP/IP networks | Ftp server and Apache web server

Programming with Java

Introduction to Java | Java Basic Syntax | Data Types and Variables | Inheritance / Polymorphism | Expressions / Operators | Functions / Classes | Exception Handling | Loops / Strings / Arrays | Interfaces and Packages | Threading | Files Input / Output | Java Applets

Hadoop: Big Data Analysis Framework

Introduction to Big Data | Operational vs. Analytical Systems | Big Data Solutions | Traditional Enterprise Approach | Google's Solution | Hadoop Architecture | Hadoop Distributed File System | Environment Setup | Downloading Hadoop | Hadoop Operation Modes, Verifying Hadoop Installation | HDFS Architecture | HDFS Operations | HDFS Command Reference | Map Reduce | The Algorithm | Inputs and Outputs (Java Perspective) | Terminology, Example Scenario | Compilation and Execution of Process Units Program | Important Commands | Introduction to Apache | Apache Pig | Apache HIVE | Apache HBASE and Apache Ecosystem | Introduction to Streaming | Multi-Node Cluster | Creating User Account | Mapping the nodes | Configuring Key Based Login | Installing and Configuring Hadoop | Starting Hadoop Services | Adding a User and SSH Access | Adding and Removing Data Node from the Hadoop Cluster.



- After completion of this course candidate can easily able understand Big Data Technologies.
- They can deal with Hadoop Distributed File System.
- They can use HDFS architecture.
- They can handle big data with Map Reduce algorithms.

Batches at Select Centres Only.