Document Generated: 08/31/2024
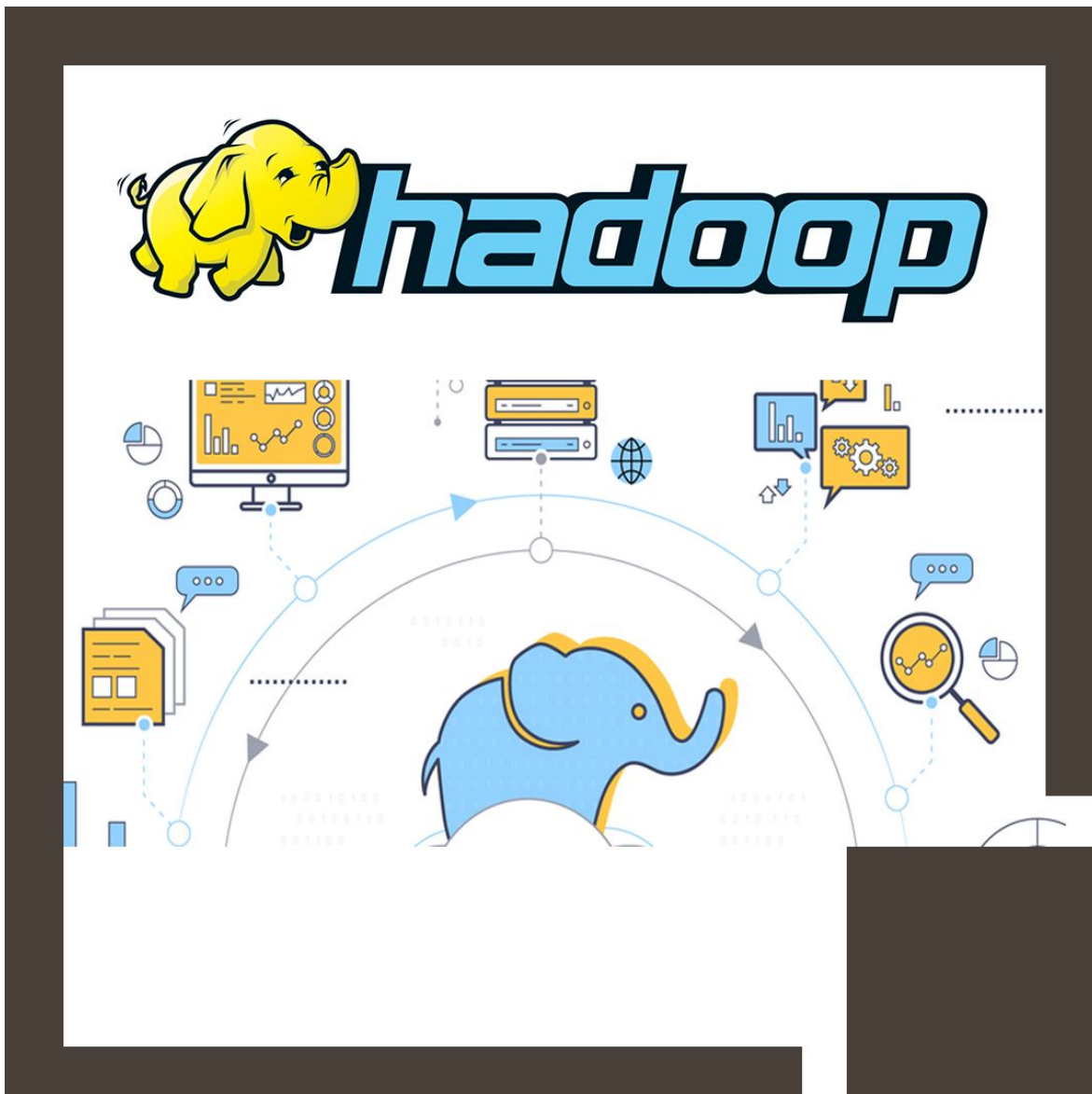
Learning Style: Virtual Classroom

Provider: Java

Difficulty: Beginner

Course Duration: 3 Days

Introduction to Hadoop Administration (TTCHADADM3)

About this course:

Apache Hadoop is a course of an open-source system utilized for creating reliable and distributing compute clusters. Hadoop is credited with the Watson Jeopardy of IBM victory of 2011 and can be utilized nearby other related systems, to process significant unstructured or semi-structured informational indexes from a few sources to the classify, gain from, analyze, and set forward proposals for choice support, business investigation, and other propelled types of machine intelligence.

The introductory-level, hands-on lab-driven course is focused on the administrators who are new to Hadoop and has been made answerable for the maintenance of Hadoop clusters and related parts. Hadoop is intended for tremendous versatility, and it is very tolerant of issues when contrasted with other cluster structures. As administrators, students should install, design and maintain Hadoop on Linux, among different computing situations.

The course motivation might be redone effectively in request to address regions of certain interest to the group. There are lab varieties that help Cloudera and Hortonworks distributions too.

The normal compensation of a Hadoop Administrator is $122,005 every year.

Course objective:

- Install, design and maintain Apache Hadoop structure
- Discover Spark, YARN, and MapReduce
- Investigate Mahout, MLib, and different systems
- Tune and upgrade the presentation of Hadoop
- Hadoop engineering
- Install Hadoop
- Test Hadoop programs
- Install Hadoop for HBase and the cloud

Audience:

Administrators who are maintaining the cluster of Hadoop and related parts in Linux conditions

Prerequisite:

No prerequisites for the course.

Course Outline:

Module 1: Hadoop Overview

- Map/Reduce

- Hadoop, YARN, and Spark
- Mahout and MLib
- Alternate Frameworks

Module 2: Hadoop Architecture

- Hadoop Map/Reduce
- YARN
- HDFS
- Spark
- Cassandra
- HBase
- Hive
- Pig

Module 3: Installing Hadoop

- Linux Considerations
- SSH Configuration
- Hadoop Installation
- OS Security
- NamedNodes
- Job Trackers

Module 4: Test-Running Hadoop Programs

- Simple MapReduce Test
- Spark Test
- Pig Test

Module 5: Cloud Installations

- Amazon EC2
- Amazon Elastic MapReduce
- Rackspace
- Installing with Docker

Module 6: Optimization and Tuning

- Performance Metrics
- Node Sizing
- Kernel Tuning

Module 7: Installing HBase

- HBase Installation
- ZooKeeper

Module 8: Previewing Hadoop 3

Credly Badge:

Display your Completion Badge And Get The Recognition You Deserve.

Add a completion and readiness badge to your Linkedin profile, Facebook page, or Twitter account to validate your professional and technical expertise. With badges issued and validated by Credly, you can:

- Let anyone verify your completion and achievement by clicking on the badge
- Display your hard work and validate your expertise
- Display each badge's details about specific skills you developed.

Badges are issued by QuickStart and verified through Credly.

Find Out More or See List Of Badges