

## Bigdata Testing

**Mode of Training:** Online, Classroom, Corporate

**Faculty:** Mr.

**Course Duration:**



**Course  
Curriculum**

## How Big data Testing different from Manual testing

In Big data testing QA engineers verify the successful processing of huge amount of data using commodity cluster and other supportive components. It demands a high level of testing skills as the processing is very fast.

Big data Testing implies – Testing the entire process using Hadoop components with the help of test cases and Rules Mapping document.

## Who are eligible to take big data Testing Training Course?

- Testers who want to get job on big data testing, to get more salary and also who want to get job quickly.
- Who are having very basic knowledge on Database concepts and manual testing are eligible to attend training.
- This training is essential for those who are seeking increments and promotions to enhance their skills & to become a professional in database management.
- There's quite a lot of demand for big data Testing resources in the market.

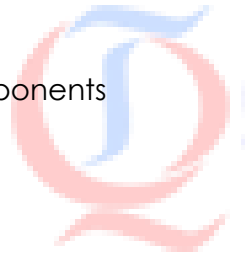
**Prepare yourself for the future and upgrade your skillet and get  
Opportunity ASAP.**

## Understanding Big Data and Hadoop:

- What is Big Data?
- 3 V's Concepts
- Diff Problems and Solutions of Big data

## Hadoop Architecture:

- What is Big Data?
- What is Hadoop and History of Hadoop?
- Hadoop Architecture
- Hadoop ecosystem components
- Hadoop Storage: HDFS
- Hadoop Processing: MapReduce Framework
- Hadoop Server Roles: Name Node Secondary Name Node, and Data Node
- Anatomy of File Write and Read.
- Different Components of Hadoop.



## Introduction to UNIX COMMANDS:

- Echo, cat, cp, cut, ls, Cd, touch
- mkdir, rmdir, rm,

- chmod, mv, head, tail
- Date, PS, uniq, diff
- grep,vi editor

## **HDFS (Hadoop Distributed File System):**

- Significance of HDFS in Hadoop
- Features of HDFS
- 5 Daemons of Hadoop
  - Name Node and its functionality
  - Data Node and its functionality
  - Resource Manager and its functionality
  - Node Manager and its functionality
  - Secondary Name Node and its functionality
  - Introduction about Blocks
  - Data Replication
  - Data storage in Data Nodes
  - Replication Configuration
  - Fail Over Mechanism
  - Replication Factor
  - Changing block size for file and Directory

- File processing in HDFS
  - How to process files in edge node?
  - How to load files into HDFS?
  - File validation between Source and Target
  - Test cases preparation for File processing

### **Apache PIG and Pig Latin:**

- Introduction to Apache PIG
- MapReduce Vs Apache PIG
- SQL Vs Apache PIG
- Physical & Logical Layer
- Different Data types in Apache PIG
- Modes of Execution in Apache PIG
  - Local Mode, Map Reduce or Distributed Mode
- Execution Mechanism
  - Grunt shell, Script, Embedded
- Transformations in PIG
- How to write a simple PIG Script
- Hands on with PIG lactic script
- Validation on PIG scripts
- Test cases preparation for PIG scripts

## Hive and Hive QL:

- HIVE Introduction
- Hive Architecture and Installation
- Comparison with Traditional Database
- Operators and Functions
- Hive Meta Store and Integration with MySQL
- Hive integration with Hadoop
- SQL vs. HIVE QL
- Partitioning, Dynamic Partitioning and Bucketing
- Hive Tables (Managed Tables and External Tables, Storage Formats, Importing Data, Altering Tables, Dropping Tables)
- Hive data format – Text, ORC, Avro, parquet

## SQOOP:

- Introduction to SQOOP
- How to connect relational database using SQOOP
- Different Sqoop Commands
  - Different flavors of imports, Export, HIVE imports
- Hands on with Examples

## Flume:

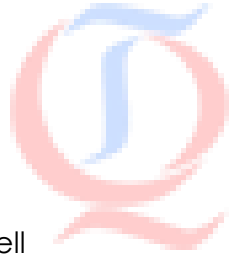
- Flume Introduction
- Flume Architecture
- Flume Master, Flume collector and Flume Agent

## Oozie:

- Oozie Introduction
- Oozie Architecture
- Oozie Configuration files
- Oozie Job Submission
- Workflow.xml
- Coordinator.xml
- Job. Coordinator. Properties

## SPARK:

- Introduction to Spark
- Spark Vs Map Reduce Processing
- Architecture of Spark
- Spark Shell Introduction
- Creating Spark Context
- File Operations in Spark Shell
- Real time Examples of Spark
- Spark Components
  - Spark Core
  - Spark SQL
  - Spark Streaming
  - Spark ML Lib



## Hadoop Testing Concepts:

- Importance of Hadoop Testing
- Responsibilities of Hadoop Tester
- Real time scenarios for Hadoop Tester

- Different approaches for Hadoop Testing
- How to validate data for multiple sources
- How to prepare Hadoop Test Plan
- How to prepare Estimates
- How to design Test Cases
- How to design Test Scenarios
- How to prepare Test scripts
- How to prepare Test data for test cases
- How to execute Hadoop Test Cases
- Defect Tracking and Reporting
- Types of defect in Hadoop Testing
- Challenges in Hadoop Testing

### **Project Essentials:**



- Project Name:
- Project Description:
- Client Name:
- Client Description:
- Business Background of Project:
- Process followed in the Project:
- Tools used in the Project:
- Environments in the Project:
- Team Size:
- Project Architecture:
- Testing Life Cycle in the Project:



- Query Tracker:
- Test Plan preparation:
- Test Cases preparation:
- Requirement Traceability Matrix:
- Review process in Project:
- Test case Execution:
- Defect Life Cycle in project:
- Role of Quality Centre in project:
- Regression Testing in Project:
- Bugs Identified in Project:
- Risks Identified in Project:
- Challenges identified in Project:

#### **Additional Services from Quality Thought:**

- Practical Sessions
- Interview KIT
- Resume Preparation
- Mock Interviews
- Support from Quality Thought to get the job and after got the job also

# Supporting Enterprises around the Globe



# Our Other Courses

## DevOps & Cloud Computing

- [AWS Admin Training](#)
- [Devops Training](#)
- [Azure Admin Training](#)

## Artificial Intelligence & IOT

- [Artificial Intelligence Training](#)
- [IOT Training](#)

## Big data

- [Hadoop Development Training](#)
- [Spark And Scala Training](#)

## Data Science & Blockchain

- [Data Science Training](#)
- [Blockchain Training](#)

## QA and Testing

- [Manual Testing](#)
- [Big data Testing Training](#)
- [Selenium Training](#)
- [Appium Training](#)
- [Web services Testing Training](#)
- [Security Testing Training](#)
- [Performance Testing Training](#)
- [TOSCA Training](#)
- [ETL Testing Training](#)

## RPA and Automation

- [Automation Anywhere Training](#)
- [Blueprism Training](#)
- [UiPath Training](#)

## Programming & Framework

- Core Python
- [Advanced Python Training](#)
- [Django Training](#)
- [C,C++Programming Training](#)

## Others courses

- [Digital Marketing Training](#)
- [Spoken English Training](#)
- [Chat Bot Making Training](#)
- [BI Reporting Tools Training](#)
- Sail Point IIQ

Established in 2010

**QualityThought**<sup>®</sup>

The Leader in Software Training



[www.qualitythought.in](http://www.qualitythought.in)



+91 996 379 9240



[info@qualitythought.in](mailto:info@qualitythought.in)