



## Course Details

This 32 hours Instructor lead training course is aimed at aspiring administrators to teach big data fundamentals and administration of Apache Hadoop cluster using Apache ecosystem tools like Ambari, Knox and Ranger. This course, part of the Big Data program at AiQuest, is designed for administrators who will be managing the Hadoop Platform (HDP) and services for the big data cluster. This course covers installation, configuration and other typical cluster maintenance tasks.

Students will learn the details of Hadoop, YARN, the Hadoop Distributed File System (HDFS), MapReduce, and dive deep into HDP concepts like – configuration and deployment, troubleshooting, High Availability (HA), Security and Operations.

This training course is the best suited for certification aspirants (HDPCA and CCA Administrator certifications) and prepares Hadoop administrators for real-world challenges.

## Prerequisites

Students should be familiar with any programming language or scripting. SQL and basic UNIX knowledge is helpful but not necessary. No prior Hadoop knowledge or experience is required.

## Course Outline

- Big Data Overview
- Introduction to Apache Hadoop
  - Hadoop Overview
  - Hadoop ecosystem projects overview
- Apache Hadoop file storage
  - HDFS overview
  - HDFS Architecture
- Apache Hadoop Data Processing framework
  - MapReduce Overview
  - MapReduce Architecture
  - YARN overview
  - YARN Architecture
  - Demo on MapReduce Jobs



- Hadoop Installation using Ambari
  - Apache Ambari
- Ambari Overview
  - Configure a local HDP repository
  - Install Ambari-server & Ambari-agent
  - Hadoop cluster Architecture overview
  - Hadoop Data platform (HDP)
  - Cluster Expansion and Maintenance
  - Introducing new services in cluster
- Hadoop Configuration
  - Rack topology – Definition and Deployment
  - Hadoop service configuration
  - User permissions
  - DataNode files configuration
- Troubleshooting
  - How to Restarting Hadoop services
  - Reading Service Logfiles
  - How to configure and manage alerts
  - Troubleshoot a failed job
- High Availability
  - Setup HDFS high availability
  - Setup YARN high availability
  - Setup Hive Server2 high availability
- Security
  - Apache Knox
    - Knox Overview
  - Apache Ranger
    - Ranger overview
    - Installation and Demonstration Lab
- Operations
  - Access control Overview and Setup
  - Snapshot overview
  - Demonstration Lab



## **Course Duration**

32 hours (16 hours theory and concepts; 16 hours practical labs and demos)

This course is spread over 4 weekends (Saturday and Sunday) 6.30 AM to 10.30 AM Eastern Standard Time (GMT- 4:00 hrs). Course could be customized to the needs of participants.

## **Cost**

Fee per participant varies based on the course delivery method and extent of customization.

The cost includes training, material and cloud-based lab fees.

Note: Please inquire with us for ongoing promotions and early bird prices.

## **Registration and Enquiries**

For registration and any other inquiries, please contact AiQuest at [connect@aiquestince.com](mailto:connect@aiquestince.com) or call us on 514-910-6785. Visit us at [www.aiquestinc.com](http://www.aiquestinc.com)

## **About AiQuest**

The mission of Ai Quest (AiQ) is to help organizations and knowledge workers to explore and realize their true potential in the Artificial Intelligence (AI) landscape. The true potential of Big Data in the AI realm goes beyond implementing new technologies and having appropriate data analytics. The strategy must include well trained resources, right performance measures that affect the corporate performance, exploiting existing technological resources to maximize the value and continuous investment in corporate training.

## **Course design philosophy**

We want to bring corporate quality and industry standard training to individuals seeking a career in Big Data. Our courses are modelled based on extensive industry experience and cater to current Industry needs to provide relevant practical experience and real-time working knowledge. Our elite courses cover core



concepts in Big data as offered by corporate solution partners-Horton Works, Cloudera and Pivotal.

For those looking to certify, the course has been designed specifically to help take the certification examination with ease. Also, the courses are designed with an ideal theory to practical ratio of 50:50, ensuring learning conceptual knowledge backed by practical applicable skills relevant to the work force. The courses are delivered by professional trainers who offer corporate trainings to companies and are working as consultants and architects on Big Data projects.

