



Hortonworks University announces availability of the Admin 2.0 exam.

Obtain your [Hortonworks Certified Apache Hadoop 2.0 Administrator](#) qualifications today. Being a leader in the Hadoop community confirms your mastery of the skills and tools needed to configure and manage Hadoop systems. A Hortonworks certification recognizes your growth as a professional and increases recognition by your peers, your employer and other leaders in the industry.

The path to certification

Prepare for the Hortonworks Certified Apache Hadoop 2.0 Administrator exam by attending the 4-day [Apache Hadoop 2.0: Operations Management with the Hortonworks Data Platform](#) course. The course, designed for administrators who deploy and manage Apache Hadoop 2.0 clusters, teaches students how to install, configure, maintain and scale the Hadoop 2.0 environment. Course topics include:

- Introduction to HDP and Hadoop 2.0
- Hadoop Distributed File System (HDFS) architecture
- HDP2.0 installation prerequisites and planning
- Configuring Hadoop using Apache Ambari
- Ensuring data integrity
- HDFS NFS gateway
- YARN architecture and MapReduce
- Job schedulers
- Enterprise data movement
- HDFS web services
- Hive architecture and administration
- Transferring data with Sqoop
- Data streaming using Flume
- Workflow management using Oozie
- Monitoring HDP 2.0 services
- Commissioning and decommissioning nodes
- Backup and recovery
- Rack-aware clusters
- NameNode High Availability (HA)
- Securing HDP

Taking the exam

The Hortonworks Certified Apache Hadoop Administrator 2.0 exam consists of 48 match and multiple-choice questions. The exam is delivered in English. A passing score is 75%. Certification candidates may take two practice exams at no charge. Register at the [certification site](#).

This exam is administered through Kryterion, Inc. The exam can be sat at authorized testing center or via remote proctoring. For additional information and to register please visit the certification site.

If a candidate does not pass an exam on the first attempt, he or she may re-register and sit the exam as soon as the final score is delivered. After the second attempt a candidate must wait 7 calendar days from their original appointment time before he or she can register to retake the exam. Should a candidate need to retake the exam again there will be a 10-days waiting period. Once the exam is passed, a candidate may not make any further attempts.

Core exam topics

1. Objective 1.1 – HDFS

- Understand HDFS architecture
- Understand how the NameNode maintains the file-system metadata
- Understand how data is stored in HDFS
- Understand the relationship between NameNodes and DataNodes
- Understand the relationship between NameNodes and namespaces in Hadoop 2.0
- Understand the WebHDFS commands
- Understand various “hadoop fs” command

2. Objective 2.1 – Installing and configuring HDP

- Understand minimum hardware and software requirements
- Understand how to setup a local repository for HDP installation
- Understand How to install HDP using Apache Ambari
- Understand difference between master and slave services
- Understand complete deployment layout
- Understand how to configure and manage different services
- Understand different configuration parameters

3. Objective 3.1 – Ensuring Data Integrity

- Understand block-scanning report
- Running file-system check
- Understand replication factor, under & over replication
- Setting up NFS Gateway for easy access to HDFS data

4. Objective 4.1 – YARN Architecture and MapReduce

- Understand the architecture of YARN
- Understand the components of the YARN ResourceManager
- Demonstrate the relationship between NodeManagers and ApplicationMasters
- Demonstrate the relationship between ResourceManagers and ApplicationMasters
- Explain The relationship between Containers and ApplicationMasters
- Explain how Container failure is handled for a YARN MapReduce job
- Understand the architecture of MapReduce
- Understand the various phases of a MapReduce job

5. Objective 5.1 – Job Schedulers and Enterprise data movement

- Understand concept of job scheduling
- Configuring The capacity scheduler
- Understand difference between capacity and fair scheduler
- Understand various data ingestion mechanism for Hadoop
- Difference between traditional and hadoop based ETL platform
- Using distcp command to move data between 2 clusters
- Understand Hive architecture
- Data movement between traditional DB and hadoop using Apache sqoop
- Hive/MR vs Hive/Tez
- Data Streaming using Apache Flume
- Workflow configuration and deployment using Apache oozie

6. Objective 6.1 – Cluster Monitoring and Administration

- Understand monitoring using Ambari UI, Ganglia and Nagios
- Commissioning and decommissioning nodes
- Understand different ways to backup and recover Hadoop data
- Using Hadoop snapshots
- Rack awareness and topology
- Understand NameNode High Availability
- Using “hdfs haadmin” commands

7. Objective 7.1 – Securing HDP

- Understand security concepts
- Kerberos configuration
- HDP authorization and authentication