HADOOP CERTIFICATIONS

The Hadoop certification landscape

The Apache Hadoop organization provides tutorials on Hadoop technology but does not offer certifications or endorse any certification programs. Hadoop distribution vendors have taken on the professional development by creating their own education, testing, and certification programs. These programs are specific to the distribution environment, but the fundamental education about core Hadoop components like HDFS, YARN, Hive, Pig, Spark, etc. is transferrable to other environments. The following sections provide an overview of the certification programs from the three leading Hadoop distribution vendors.

Cloudera

Cloudera University provides certification training and other courses in tracks for administrators, data analysts, and developers. Certifications currently available through the Cloudera Certified Professional (CCP) Program include:

- CCP Data Engineer: Focuses on how to build the "pipelines" to produce data sets that are optimized for different types of workloads.
- CCP Data Scientist: Teaches how to develop production-ready, scalable solutions.

Cloudera also offers exams for Cloudera Certified Associate (CCA) which provide the technical knowledge foundation for the Cloudera Hadoop environment.

- CA Spark and Hadoop Developer: To earn this certification individuals must demonstrate competency in writing Python and Scala code to run on Hadoop clusters.
- Cloudera Certified Administrator for Apache Hadoop (CCAH): Certification of competency in configuring, deploying, securing, and maintaining Hadoop clusters.

Hortonworks

Hortonworks University is the company's training and certification program. Some of the certification programs offered by Hortonworks include:

- **HDP Certified Developer (HDPCD):** for Hadoop developers using frameworks like Pig, Hive, Sqoop and Flume.
- **HDP Certified Administrator (HDPCA):** for administrators who deploy and manage Hadoop clusters.
- **HDP Certified Java Developer (HDPCD:Java):** for developers who design, develop and architect Hadoop-based solutions written in the Java programming language.

In addition to the certification programs, Hortonworks offers self-paced and classroom training designed for developers, systems administrators, and data analysts.

MapR

MapR Academy is MapR's training and certification program. It currently offers four different certifications:

- **MCHA - MapR Certified Hadoop Administrator:** This certification demonstrates proficiency in the administration of Hadoop Clusters and MapR administration tools.

- **MCHD - MapR Certified Hadoop Developer:** This certification exam demonstrates proficiency in the design and development MapReduce/YARN programs.
- **MCHBD - MapR Certified HBase Developer:** This certification reflects proficiency in the HBase development programs that use HBase as a distributed NoSQL datastore.
- **MCSD - MapR Certified Spark Developer:** This certification demonstrates proficiency in Apache Spark programming to work with large datasets.

MapR Academy also offers numerous on-demand courses and other resources in addition to the certification courses.

Note there are other big data training and certification programs that are not Hadoop specific. For example, Microsoft, Oracle, and SAS Institute are among the technology companies that have their own big data education programs.