

Section A: Multiple Choice Questions (1 marks each)

10 Marks

Q.1 How many versions Hadoop has

- a) One
- b) Two
- c) Three
- d) Four

Q.2 A Data node Machine contains exactly

- a) One Resource Manager
- b) One Application Master
- c) One Node Manager
- d) One MapReduce job

Q.3 Which one of the following statements is true regarding a MapReduce job?

- a) Input to mapper is sorted key, value pairs
- b) Number of mapper tasks is equal to number reducer tasks.
- c) Reducer starts only when all mapper finish execution
- d) All of the above

Q.4 Which of the following statements is correct for an external table of hives?

- a) Source Data gets deleted if external table is dropped
- b) Source Data does not get deleted if the external table is dropped.
- c) It provides transaction capabilities like RDBMS
- d) None of the above

Q.5 which is the default hive metastore

- a) Derby
- b) MySQL
- c) Spark-SQL
- d) HBase

Q.6 Which among these can act as hive execution engine

- a) MapReduce
- b) Spark
- c) TEZ
- d) Any of the above

Q.7 MongoDB can be used as backend database in which of following use cases?

- a) Video gaming application
- b) Content management application
- c) E-commerce applications
- d) All of the above

Q.8 A collection in mongoDB is

- a) a group of related documents
- b) A group of data bases
- c) A group of schema
- d) A group of rows

Q.9 In Spark, RDD stands for

- a) Rigid Distributed Dataset
- b) Resilient Distributed Dataset
- c) Realtime Distributed Dataset
- d) Random distributed Dataset

Q.10 The various ways in which a Spark RDD can be persisted are

- a) Memory only
- b) Memory only serialized
- c) Memory and hard drive
- d) All of the above

Section B: Multiple Choice Questions (2 marks each)

10 Marks

Q.1 In Hive, command to change name of a table from t1 to t2 is :

- a) UPDATE TABLENAME FROM t1 to t2;
- b) ALTER TABLE t1 RENAME TO t2;
- c) RENAME TABLE FROM t1 to t2;
- d) UPDATE TABLE t1 AS t2;

Q.2 In MongoDB, Select the correct query for displaying users name based on ascending order of the age

- a) db.users.find().sort({age:1})
- b) db.users.find().limit()
- c) db.users.find()
- d) db.users.find().sort(age:-1)

Q.3 which is the correct command syntax to copy file f1 from local directory to HDFS directory

- a) hadoop fs -put <local_directory_path> <hdfs_directory_path>
- b) hadoop fs -copyFromLocal <hdfs_directory_path> <local_directory_path>
- c) hadoop fs -copy <local_directory_path> <hdfs_directory_path>
- d) hadoop fs -move <hdfs_directory_path> <local_directory_path>

Q.4 How to rename the column name using Pyspark in Spark DataFrame

- a) df.withColumnRenamed(existingName, newName)
- b) df.withRenamed(existingName, newName)
- c) df.withColumn(existingName, newName)
- d) df.Renamed(existingName, newName)

Q.5 How to convert RDD into DataFrame using Pyspark

- a) spark.createDataFrame(rdd, schema = Columns)
- b) rdd.toDF(columns)
- c) spark.convertDataFrame(rdd, schema = Columns)
- d) A & B

e) A & C

f) B & C