



NBV-003-1082002

Seat No. 20240
108200

M. Sc. (I.T.) (Sem. II) (CBCS) Examination

April / May - 2017

CS - 08 : Introduction to Big Data & Hadoop

Faculty Code : 003

Subject Code : 1082002

Time : $2\frac{1}{2}$ Hours]

[Total Marks : 70

- 1 (A) Answer the following : 4
- (1) _____ command is used to move file from source to destination
 - (2) On startup, the NameNode enters a special state called Safemode. True or False?
 - (3) Give full form of CRC.
 - (4) Give full form of HDFS.
- (B) Answer any **one** in brief : 2
- (1) What is virtual box and virtual machine?
 - (2) Explain rack awareness and block placement policy.
- (C) Answer any **one** in detail : 3
- (1) Explain CAP theorem.
 - (2) List and explain advanced HDFS features.
- (D) Answer any **one** : 5
- (1) Explain HDFS architecture.
 - (2) Explain HDFS data flow (Read vs. Write).
- 2 (A) Answer the following : 4
- (1) A _____ node acts as the Slave and is responsible for executing a Task assigned to it by the Job Tracker.
 - (2) Reduce() function is responsible for consolidating the results produced by each of the Map() functions/ tasks. True or False?

- (3) Shuffled and sorted data is going to pass as input to the reducer. True or False?
- (4) Job scheduling, resource management and job monitoring are being done by _____ in hadoop.
- (B) Answer any **one** in brief : 2
- (1) Explain macros.
- (2) What is slot and container?
- (C) Answer any **one** in detail : 3
- (1) Explain Load, Store and Dump with suitable example.
- (2) Explain Data flow of MapReduce.
- (D) Answer any **one** : 5
- (1) Explain MapReduce Architecture.
- (2) Explain architecture of Pig.
- 3 (A) Answer the following : 4
- (1) _____ implementation classes can be used as values in Map Reduce programming,
- (2) Tables or partitions are sub-divided into _____ in hive.
- (3) Writable Comparable classes can be used as keys in MapReduce programming. True or False?
- (4) _____ language is used to execute Hive statements.
- (B) Answer any **one** in brief : 2
- (1) Explain select statement with suitable example.
- (2) How to load data in hive?
- (C) Answer any **one** in detail : 3
- (1) Explain sorting algorithm.
- (2) Explain external and managed table.

- (D) Answer any **one** : 5
- (1) Explain architecture of Hive.
 - (2) List optimization techniques and explain any one.
- 4 (A) Answer the following : 4
- (1) _____ is used to delete table in HBase.
 - (2) _____ is used to create table in HBase.
 - (3) HBase is a distributed _____ database built on top of the Hadoop file system.
 - (4) The Master Server assigns regions to the region servers and takes the help of Apache Zookeeper for this task. True or False?
- (B) Answer any **one** in brief : 2
- (1) Explain use of Filter in HBase.
 - (2) Explain Hbase Data Modeling.
- (C) Answer any **one** in detail : 3
- (1) How to drop a table in HBase? Explain it with example.
 - (2) Explain catalog tables.
- (D) Answer any **one** : 5
- (1) Explain HBase coprocessors.
 - (2) Write a note on HBase Architecture.
- 5 (A) Answer the following : 4
- (1) _____ is an open-source cluster-computing framework designed for fast computation.
 - (2) _____ is a scheduler system to run and manage Hadoop jobs in a distributed environment.
 - (3) Give full form of RDD.
 - (4) Give full form of EMR.

- (B) Answer any **one** in brief : 2
- (1) How to export data using Sqoop?
 - (2) What is graph processing?
- (C) Answer any **one** in detail : 3
- (1) How to create workflow in Oozie?
 - (2) How Flume works?
- (D) Answer any **one** : 5
- (1) Explain Spark Architecture.
 - (2) Write a note on Neo4j.
-