Big Data and Hadoop - 01

1. What is the default replication factor of Hadoop cluster?

a. 3

b. 2

c. 4

d. 1

ANS: a

2. Which component in Hadoop Cluster is responsible for serving read and write requests from

the file system's clients?

a. Name Node

b. Data Node

c. Both a & b

d. None of the above

ANS: b

3. Which component of Hadoop Cluster manages the file system namespace and regulates access

to files by clients?

a. Name Node

b. Data Node

c. Both a & b

d. None of the above

ANS: c

4. If a file size of size 100 MB is stored on HDFS, what would be the split size?

a. 64 MB & 64 MB

b. 64 MB & 36 MB

c. 100 MB

d. None of the above

ANS: d

5. State true or false: MR2 support various MPP modes for data processing?

a. FALSE

b. True

ANS: b

6. Which comand of HDFS helps copy files from HDFS to Local file system?

a. copyFromLocal

b. copyToLocal

c. put

d. mv

ANS: c

7. Which Eco system component of Hadoop is good for non sql programmers?

a. Hive

b. Hbase

c. Flume

d. Pig

ANS: a

8. Block size of a Hadoop cluster is configurable by Administrator?

a. TRUE

b. FALSE

ANS: a

9. The functions performed by DataNodes in Hadoop Cluster is/are?

a. Data Block Creation

b. Data Block Deletion

c. Data Block Replication

d. All above

ANS: d

10. Find error in below command:

hdfs dfs -put /home/user1/abc.txt

a. Target name missing

b. Source name should include hdfs://

c. No error

ANS: c

11. Hadoop block size should be multiple of which unit?

a. 32 MB

b. 50 MB

c. 64 MB

d. 70 MB

12. Which component of the hadoop cluster manages data on slave nodes?

a. Name node

b. Data Node

c. Task Tracker

d. Job Tracker

ANS: c

13. MR1 and MR2 are two modes of processing in Hadoop?

a. TRUE

b. FALSE

ANS: a

14. What is Hadoop?

a. Open source software for reliable, scalable, distributed computing.

b. A framework that allows for the distributed processing of large data sets across clusters of computers using simple programming models.

c. Both a & b

d. None of the above

ANS: c

15. Hadoop provides

a. A reliable distributed storage and processing system

b. Only distributed storage

c. Only processing system

d. None of the abov

ANS: d