

Week 2

HDFS

1. When traditional file systems can handle file sizes in exabytes what is the need for HDFS?
 - a. Traditional file systems are not reliable
 - b. Traditional file systems are not suitable for distributed computing

Answer: b

2. What is the default replication factor of a block on HDFS?
 - a. 1
 - b. 2
 - c. 3

Answer: c

3. What is the HDFS commands to check the blocks and the block locations?
 - a. Chmod
 - b. Fsck
 - c. copyToLocal

Answer: b

4. Namenode stores the block locations in its hard disk
 - a. True
 - b. False

Answer: b

5. How many data can be stored at one cluster assuming the cluster comprises for 120 racks with 2 data node in each rack with data node size as 64GB and the file size as 64MB?
 - a. 81920
 - b. 80250
 - c. 1024

Answer: a

6. Each Cluster has a number of data stores called as
 - a. Nodes
 - b. Racks
 - c. Blocks

Answer: b

7. The masters have a different configuration supporting high DRAM and processing power.
 - a. True
 - b. False

Answer: a

8. The NameNode stores all the file system related information such as
- a. The file section is stored in which part of the cluster
 - b. Last access time for the files
 - c. User permissions like which user has access to the file.
 - d. All

Answer: d

9. Hadoop is write once
- a. True
 - b. False

Answer: a

10. What is the best hardware configuration to run Hadoop?
- a. dual core machines or dual processors with 4GB or 8GB RAM
 - b. dual core machines or dual processors with 4GB or 4GB RAM
 - c. dual core machines or dual processors with 4GB or 16GB RAM

Answer: a

11. This is the node that keeps track of the latest checkpoint in a directory that has same structure as that of NameNode's directory.
- a. Name Node
 - b. Checkpoint Node
 - c. BackupNode

Answer: b

12. What is Commodity Hardware?
13. Can a Namenode exist with no data
- a. True
 - b. False

Answer:b

14. Whenever a client submits a Hadoop Job who receives it?
- a. Name node
 - b. Data node
 - c. Secondary node

Answer:a

15. What are the different type of configurations that Hadoop can be configured?
- a. Stand alone mode
 - b. Pseudo distributed mode
 - c. Fully distributed mode