HADOOP-HDFS

Short Questions

1. What happens when two clients try to access the same file in the HDFS?
2. How is HDFS fault tolerant?
3. What are the two types of metadata that a Name Node server holds?
4. If you have an input file of 355 MB, how many input splits would HDFS create and what would be the size of each input split?
5. Can standby name node replace the primary name node in HDFS? Justify your answer.
6. How does rack awareness work in HDFS?
7. What would happen if you store too many small files in a cluster on HDFS?
8. Who divides the file into Block while storing inside hdfs in hadoop?
9. Is concurrent write into HDFS file possible? Justify your answer.
10. Distinguish among secondary name node and standby name node.

Long Questions

1. Explain how HDFS is different than Google File System (GFS) with a neat diagram.

2. Suppose one client wants to write a file (named as abc.txt) having size 135 MB into HDFS (HADOOP 1.0). Explain how rack awareness works in HDFS for this write operation with a neat diagram.

3. Explain the following term:

    a) Secondary name node
    b) Heart beat mechanism
    c) Block
    d) Job Tracker
    e) Task Tracker

4. What do you mean by a Hadoop Distributed File System (HDFS)? Explain the concepts of Replication and Fault Tolerance in HDFS with a neat diagram.