1. HDFS is built around the idea that data is written \_\_\_\_\_but read many times. a) many b) twice c) data already exists d) once

2. Hadoop divides input into fixed size pieces called what? a) output result b) input splits c) input data d) input blogs

3. All the blocks are replicated in other nodes for \_\_\_\_\_\_ a) security b) big data c) pool d) fault tolerance

4. Block size can be changed using the properties in \_\_\_\_\_\_ a) core-site.xml b) Hadoop-env.sh c) hdfs-site.xml d) yarn-site.xml

5. Hadoop uses the \_\_\_\_\_\_representation of the data stored in the file blocks known as Input splits. a) physical b) logical c) mechanical d) none

6. DFS calls NameNode to create file in file system’s\_\_\_\_\_ a) dataspace b) resourcespace c) namespace d) nodespace

7. Data packets are streamed to first DataNode in the \_\_\_\_\_\_\_\_ a) handshake b) pipeline c) hard disk d) hdfs

8. The client has finished writing data, it calls \_\_\_\_\_\_\_on the stream. a) close() b) read() c) open() d) check()

9. Blocks are read in order, with the \_\_\_\_\_\_\_\_\_ opening new connections to datanodes as the client reads through the stream. a) DFSoutputstream b) DFSInputStream c) DFStrackManager d) DFSStringConcatination

10. If I have 100 input splits, how many maps will run? a) 200 b) 50 c) 100 d) 1