Session 3: HDFS Internals

Assignment 2

1. **Which is not the property of hdfs-site.xml file?** 
   1. Block size
   2. Replication factor
   3. Secondary NameNode port address
   4. Application server Ans

1. **What are the properties that we can edit in hdfs-site.xml file?**

a. Block size

* 1. Replication factor
  2. Block reporting interval
  3. All the above Ans

1. **Underlying storage layers where MapReduce programs are written:**

a. Abstracted Ans

b. Open to all

c. Locked

d. None of these

**4. Which is not a valid method in FSDATA Output Stream?**

a. close( )

1. open( ) Ans
2. getpos( )
3. sync( )
4. **How many blocks of size 128MB will be allocated for a file of size 524288KB:** 
   1. 2
   2. 3
   3. 4 Ans
   4. 5
5. **MapReduce Job client calculates the input split by:** 
   1. Figuring the first and last whole records in the block Ans
   2. Figuring only the first block
   3. Figuring only the last block
   4. None of these
6. **Mappers are directly related to:** 
   1. Input data
   2. Input splits Ans
   3. Output data

c. Output splits

1. **After data is written in an HDFS file, HDFS does not provide any guarantee that data are visible to a new reader until the file is \_\_\_\_\_\_**

a. closed Ans

* 1. opened
  2. writing
  3. reading

1. **What mechanism does Hadoop use to make namenode resilient to failure?** 
   1. Take backup of filesystem metadata to a local disk and a remote NFS mount Ans
   2. Store the filesystem metadata in cloud
   3. Use a machine with at least 12 CPUs
   4. Using expensive and reliable hardware

1. **All the files in a directory in HDFS can be merged together using:** a. getmerge Ans
   1. putmerge
   2. remerge
   3. mergeall