Session 3

# 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

Answer – d. Application server

1. What are the properties that we can edit in hdfs-site.xml file?
   1. Block size
   2. Replication factor
   3. Block reporting interval
   4. All the above

Answer – All the above

1. Underlying storage layers where MapReduce programs are written:
   1. Abstracted
   2. Open to all
   3. Locked
   4. None of these

Answer – a. Abstracted

1. Which is not a valid method in FSDATA Output Stream?
   1. Close()
   2. Open()
   3. Getpos()
   4. Sync()

Answer – b. Open

1. How many blocks of size 128MB will be allocated for a file of size 524288KB:
   1. 2
   2. 3
   3. 4
   4. 5

Answer – c. 4

1. MapReduce Job client calculates the input split by:
   1. Figuring the first and last whole records in the block
   2. Figuring only the first block
   3. Figuring only the last block
   4. None of these

Answer – a. Figuring the first and last whole records in the block

1. Mappers are directly related to:
   1. Input data
   2. Input splits
   3. Output data
   4. Output splits

Answer – b. Input 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 \_\_\_\_\_\_
   1. Closed
   2. Opened
   3. Writing
   4. Reading

Answer – Closed

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
   2. Store the filesystem metadata in cloud
   3. Use a machine with at least 12 CPUs
   4. Using expensive and reliable hardware
   5. Answer – a. Take backup of filesystem metadata to a local disk and a remote NFS mount
2. All the files in a directory in HDFS can be merged together using:
   1. Getmerge
   2. Putmerge
   3. Remerge
   4. Mergeall

Answer – a. Getmerge