1. What is Default replication factor and how will you change it at file level? Ans:Default replication factor is 3 and can change replication factor by configure hdfs-site.xml file

2. Why do we need replication factor > 1 in production Hadoop cluster? Ans: for fault tolerance

3. How will you combine the 4 part-r files of a mapreduce job?

4. What are the Compression techniques in HDFS and which is the best one and why? The compression techniques are bzip,snappy,gzip,LZO. You can use Bzip2 codec space priority is higher and the data will be rarely needed to be queried

5. How will you view the compressed files via HDFS command?

6. What is Secondary Namenode and its Functionalities? why do we need it? Ans:Secondary name node is a helper namenode to primary name node.it operates two things fsimage and editlogs, where fs image takes snapshots of datanode and editlogs tracks the activities in hdfs. The main purpose is while secondary name node keeps track on primary name node called as checkpoint. It reads editlogs from name node and apllies to fsimage for speed process.

7. What is Backup node and how is it different from Secondary namenode? Ans: Baxckup node is keeps and maintains in-memory and uptodate copy of fileimage and also accepts and applies online streaming in transaction.

8. What is FSimage and editlogs and how they are related? Ans:FSimage and editlogs are the part of the secondary name node and they are related for updating the metdata in secondary name node from primary name node.

9. what is default block size in HDFS? and why is it so large? Ans: the default block size in hdfs is 128 mb, it is so large because of to

maintain particular capacity in block for not searching again fro storage and reduces the metadata information.

10. How will you copy a large file of 50GB into HDFS in parllel Ans:using Distcp

11. what is Balancing in HDFS? Balancer is a tool, which makes to maintain balance data across all datanodes in cluster.

12. What is expunge in HDFS ? Ans: expunge is used to clean data in trash.