Written Assignment: Hadoop

1) List the use and responsibilities of following tools of Hadoop eco-system

   a. HDFS

   b. YARN

   c. FLUME

   d. SQOOP

   e. ZOOKEEPER

   f. HIVE

2) Explain the responsibilities of Name node and Data nodes in HDFS. Differentiate between normal file system with HDFS.

3) With diagram show how write operation is performed in HDFS. List the steps to recover from failure during HDFS write operation.

4) Explain how HDFS ensures high availability of data and services.

5) With diagram show how read operation is performed in HDFS. Explain the steps involved in read operation.

6) Explain the terms heart-beat, edit-log and check-points.

7) With neat diagram, explain the architecture of HDFS.

8) List the major components of YARN. Explain the responsibilities of YARN.

9) List out the responsibilities of Resource Manager in YARN.

10) Explain different schedulers used in YARN.

11) Explain different types of failures in Map-Reduce jobs.

12) Give different reasons for task failures in Map-Reduce job. Give the steps for recovering from Application master failures.

13) With neat diagram, show how map reduce jobs are executed in Hadoop environment.

14) Give the responsibilities of

   a. Application Master

   b. Application Manager

   c. Node Manager

   d. Resource Manager

15) With example, explain Hadoop MapReduce process.