



- Notes :
1. All questions are compulsory and carry equal marks.
 2. Draw neat and labelled diagram and use supporting data wherever necessary.
 3. Avoid vague answer and write specific points/answer related to question.

1. Either

- a) What is Big data? Explain Importance of it. **8**
- b) Explain moving Data in and out in Hadoop. **8**

OR

- c) Explain different applications of Big Data? **8**
- d) Explain distributed file System in Big Data. **8**

2. Either

- a) Explain Hadoop Architecture in detail **8**
- b) Explain the anatomy of file read and write in Hadoop. **8**

OR

- c) Explain **any four** Hadoop shell command in detail **8**
- d) What are the configuration files in Hadoop? Explain. **8**

3. Either

- a) Explain Hadoop ecosystem components in detail. **8**
- b) Explain fair scheduler and capacity scheduler in detail. **8**

OR

- c) What is Hadoop 2.0 high Availability? **8**
- d) Explain HDFS Administering in detail. **8**

4. Either

- a) Explain Architecture and Installation of HIVE in detail. **8**
- b) What is the purpose of HIVE in Hadoop? Explain. **8**

OR

- c) Why does hbase Use zookeeper? Explain. **8**
- d) Explain the Joins in HIVEQL in detail. **8**

5. Solve all question.

- a) What is Data serialization? **4**
- b) Explain Task trackers in detail. **4**
- c) What is Apache Hadoop Yarn? **4**
- d) What is Hive QL? Explain in detail. **4**
