

B.TECH
(SEM VI) THEORY EXAMINATION 2022-23
BIG DATA

Time: 3 Hours**Total Marks: 100****Note:** Attempt all Sections. If require any missing data; then choose suitably.

SECTION A

1. Attempt all questions in brief.**2 x 10 = 20**

- (a) List any five big data platforms.
- (b) Discuss the importance of hadoop technology in big data analytics.
- (c) Explain three benefits of MapReduce.
- (d) Define heartbeat in HDFS.
- (e) Define data replication in Hadoop distributed file system.
- (f) Differentiate between flume and Sqoop.
- (g) Compare and Contrast No SQL relational databases.
- (h) Explain briefly about the schedulers.
- (i) Differentiate between Pig and MapReduce.
- (j) Discuss meta store in HIVE in brief.

SECTION B

2. Attempt any three of the following:**10x3=30**

- (a) Explain Hadoop ecosystem in detail.
- (b) Discuss Master Slave and Peer-Peer replication in detail.
- (c) Examine the process of reading and writing data in HDFS by a client.
- (d) Explain how CRUD operations with example are performed in MongoDB.
- (e) Draw and explain the detailed architecture of HIVE.

SECTION C

3. Attempt any one part of the following:**10x1=10**

- (a) Detail about the analysis vs. reporting while introducing the Big Data
- (b) Elaborate various components of Big Data architecture.

4. Attempt any one part of the following:**10x1=10**

- (a) Discuss the detailed architecture of Map-Reduce
- (b) Discuss the detailed architecture of YARN along with its components.

5. Attempt any one part of the following:**10x1=10**

- (a) Demonstrate the design of HDFS and concept in detail.
- (b) Discuss in brief about the cluster specification. Describe how to setting up a Hadoop Cluster?

6. Attempt any *one* part of the following: 10x1=10

- (a) State features of Apache Spark and also explain three ways of how Spark can be built with Hadoop components.
- (b) State difference between Java and Scala. Also explain various features of Scala.

7. Attempt any *one* part of the following: 10x1=10

- (a) Explain the architecture of HIVE. Also explain data flow in HIVE.
- (b) Compare and Contrast
 - (i) Apache Pig vs Map-Reduce
 - (ii) Pig vs SQL
 - (iii) Pig vs HIVE