PAPER ID-420322

Roll No:

BTECH
(SEM VIII) THEORY EXAMINATION 2021-22
BIG DATA

Time: 3 Hours   *Total Marks: 100*

Note: Attempt all Sections. If you require any missing data, then choose suitably.

SECTION A

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

Q.no	Questions	Marks	CO
(a)	What do you mean by Apache Hadoop?	2	2
(b)	Define the term Big Data.	2	1
(c)	Why do we need Hadoop?	2	2
(d)	What do you mean by Digital data?	2	1
(e)	Define data replication in Hadoop Distributed File System.	2	3
(f)	What do you mean by serialization in Hadoop Distributed File System?	2	3
(g)	Write down briefly about the schedulers.	2	4
(h)	Briefly define NameNode under Hadoop Eco System.	2	4
(i)	What is Zookeeper?	2	5
(j)	Mention the Execution Modes of Pig.	2	5

SECTION B

2. **Attempt any *three* of the following:**   **10*3 = 30**

Q.no	Questions	Marks	CO
(a)	What are views in HIVE? What is the difference between internal and external tables in HIVE?	10	5
(b)	Describe Map Reduce framework and Map Reduce working.	10	2
(c)	Describe Structured, Semi structured and Unstructured data? Explain with example.	10	3
(d)	Describe about the "Shuffle & Sort" phase and "Reducer Phase" in MapReduce.	10	4
(e)	What are the benefits of Big Data? Discuss in detail about the 5V's characteristics of Big Data.	10	1

SECTION C

3. **Attempt any *one* part of the following:**   **10*1 = 10**

Q.no	Questions	Marks	CO
(a)	Write short note on any two of the following. i) Hadoop Eco System Frameworks. ii) Joins and subqueries. iii) IBM Big Data strategy.	10	5
(b)	Write down the statement for developing a Map Reduce application.	10	2

Roll No:

BTECH
(SEM VIII) THEORY EXAMINATION 2021-22
BIG DATA

4. **Attempt any *one* part of the following:** 10 *1 = 10

Q.no	Questions	Marks	CO
(a)	Explain the analytic processes and tools under Introduction to Big Data.	10	1
(b)	Mention about the Cluster specification? Describe how to Setting up a Hadoop Cluster?	10	3

5. **Attempt any *one* part of the following:** 10*1 = 10

Q.no	Questions	Marks	CO
(a)	Discuss Master Slave and Peer-Peer replication in detail.	10	4
(b)	Explain the Hbase concepts. How Zookeeper helps in monitoring a Cluster?	10	5

6. **Attempt any *one* part of the following:** 10*1 = 10

Q.no	Questions	Marks	CO
(a)	Give the Anatomy of a Map Reduce job run in Map Reduce.	10	2
(b)	Give the detailed description about the Analysis vs. reporting while Introducing the Big Data.	10	1

7. **Attempt any *one* part of the following:** 10*1 = 10

Q.no	Questions	Marks	CO
(a)	Describe the Compression and Serialization under the Hadoop I/O Operation.	10	3
(b)	Explain about the storage mechanism in HBase. Write a query to create table in HBase.	10	4