



Paper id: 251013

Roll No:

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

BTECH
(SEM VI) THEORY EXAMINATION 2024-25
BIG DATA

TIME: 3 HRS

M.MARKS: 70

Note: Attempt all Sections. In case of any missing data; choose suitably.

SECTION A

1. Attempt all questions in brief. 02 x 7 = 14

Q no.	Question
a.	List any five big data platforms.
b.	How does Apache Hadoop help process the data?
c.	What is HDFS in the context of Hadoop?
d.	How does Hive differ from Pig in the Hadoop ecosystem?
e.	What are the main components of the Hadoop ecosystem?
f.	What are the main components of the Hadoop ecosystem?
g.	What is the role of YARN in Hadoop?

SECTION B

2. Attempt any three of the following: 07 x 3 = 21

a.	Discuss the various types of analytics used in Big Data.
b.	Discuss Master Slave and Peer-Peer replication in detail.
c.	Discuss the detailed architecture of Map-Reduce.
d.	Discuss in brief about the cluster specification. Describe how to setting up a Hadoop Cluster?
e.	Explain the architecture of HIVE. Also explain data flow in HIVE.

SECTION C

3. Attempt any one part of the following: 07 x 1 = 07

a.	Discuss the 5 Vs used in Big Data with the help of an example.
b.	Detail about the analysis vs. reporting while introducing the Big Data.

4. Attempt any one part of the following: 07 x 1 = 07

a.	Discuss the detailed architecture of YARN along with its components.
b.	Explain the key components of the Hadoop ecosystem and describe how they work together to enable large-scale data storage, processing, and analysis.

5. Attempt any one part of the following: 07 x 1 = 07

a.	Demonstrate the design of HDFS and concept in detail.
b.	Illustrate the process of reading a block in Hadoop.

6. Attempt any one part of the following: 07 x 1 = 07

a.	Explain the components of features and components of Apache Spark.
b.	Discuss the advantages and challenges of deploying Hadoop in a cloud environment. Explain how cloud platforms enhance the scalability, flexibility, and cost-efficiency of Hadoop clusters.

7. Attempt any one part of the following: 07 x 1 = 07

a.	Explain the role of Pig Latin as a high-level scripting language for processing large datasets in cloud environments. How does using Pig Latin on cloud-based Hadoop platforms simplify data transformation and analysis compared to traditional MapReduce programming?
----	---



Paper id: 251013

Printed Page: 2 of 2
Subject Code: BCS061

Roll No:

--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

BTECH
(SEM VI) THEORY EXAMINATION 2024-25
BIG DATA

TIME: 3 HRS

M.MARKS: 70

b.	Compare and Contrast
(i)	Apache Pig vs Map-Reduce
(ii)	Pig vs SQL
(iii)	Pig vs HIVE

QP25EP1_143
/ 13-Jun-2025 1:32:29 PM | 122.185.51.242