



Paper id: 251014

Printed Page: 1 of 2  
Subject Code: BCS061H

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.	What are the 5 Vs of Big Data? बिग डेटा के 5 Vs क्या हैं?
b.	State any four characteristics of Big Data. बिग डेटा की कोई चार विशेषताएँ बताइए।
c.	What are four key privacy concerns in Big Data? बिग डेटा में चार प्रमुख गोपनीयता संबंधी चिंताएँ क्या हैं?
d.	What is the role of the JobTracker in Hadoop 1.x? Hadoop 1.x में JobTracker की क्या भूमिका होती है?
e.	What is MRUnit used for? MR Unit का उपयोग किसलिए किया जाता है?
f.	What is the default block size in HDFS and why is it significant? HDFS में डिफ़ॉल्ट ब्लॉक साइज़ क्या है और यह क्यों महत्वपूर्ण है?
g.	Define NoSQL and give two examples of NoSQL databases. NoSQL को परिभाषित कीजिए और NoSQL डेटा बेस के दो उदाहरण दीजिए।

## SECTION B

2. Attempt any three of the following:

07 x 3 = 21

a.	Explain the challenges faced by conventional systems in handling Big Data. बिग डेटा को संसाधित करने में पारंपरिक प्रणालियों द्वारा सामना की जाने वाली चुनौतियों को समझाइए
b.	Define the Map and Reduce functions in MapReduce. What are the purpose of the Shuffle and Sort phase in Map Reduce? Map Reduce में Map और Reduce फ़ंक्शन को परिभाषित कीजिए। Map Reduce में Shuffle और Sort चरण का क्या उद्देश्य होता है?
c.	Describe the architecture and working of YARN. How does it improve resource management in Hadoop 2.0? YARN की संरचना और कार्य प्रणाली का वर्णन कीजिए। यह Hadoop 2.0 में संसाधन प्रबंधन को कैसे बेहतर बनाता है?
d.	Mention five vital differences between Hive and traditional RDBMS. Hive और पारंपरिक RDBMS के बीच पाँच महत्वपूर्ण अंतर बताइए।
e.	Discuss Master Slave and Peer-Peer replication in detail. मास्टर-स्लेव और पीयर-टू-पीयर रेप्लिकेशन पर विस्तृत रूप से चर्चा कीजिए।

## SECTION C

3. Attempt any one part of the following:

07 x 1 = 07

a.	Describe the architecture and key features of the Hadoop Distributed File System (HDFS). Hadoop वितरित फ़ाइल प्रणाली (HDFS) की संरचना और प्रमुख विशेषताओं का वर्णन कीजिए।
b.	Discuss in brief about the cluster specification. Describe how to setting up a Hadoop



Paper id: 251014

Printed Page: 2 of 2  
Subject Code: BCS061H

Roll No:

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

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

TIME: 3 HRS

M.MARKS: 70

	Cluster? क्लस्टर स्पेसिफिकेशन के बारे में संक्षेप में चर्चा कीजिए। Hadoop क्लस्टर को सेटअप करने की प्रक्रिया का वर्णन कीजिए।
--	---

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

a.	Discuss how Hadoop ensures security and how administrators can monitor and maintain an HDFS environment. चर्चा कीजिए कि Hadoop सुरक्षा सुनिश्चित कैसे करता है और प्रशासक HDFS वातावरण की निगरानी और रख-रखाव कैसे कर सकते हैं।
----	--

b.	Differentiate between flume and Sqoop in detail. Flume और Sqoop के बीच विस्तार से अंतर बताइए।
----	--

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

a.	Describe the role of the Hadoop Ecosystem Hadoop. Hadoop इकोसिस्टम की भूमिका का वर्णन करें।
----	--

b.	Explain the architecture of Apache Hive. How does it process and execute queries using HiveQL? Apache Hive की संरचना का वर्णन कीजिए। यह HiveQL का उपयोग करके क्वेरीज़ को कैसे प्रोसेस और निष्पादित करता है?
----	--

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

a.	Discuss Mongo DB CRUD operations (Create, Read, Update, Delete) with syntax examples. MongoDB CRUD ऑपरेशन्स (Create, Read, Update, Delete) पर चर्चा कीजिए और सिंटैक्स उदाहरण के साथ समझाइए।
----	--

b.	Compare and contrast Hive, Pig, and HBase in terms of data model, query language, and use case. डेटा मॉडल, query language, और use case के दृष्टि कोण से Hive, Pig, और HBase का तुलना और अंतर कीजिए।
----	--

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

a.	Describe modern data analytic tools and their advantages in handling Big Data. आधुनिक डेटा विश्लेषण उपकरणों का वर्णन कीजिए और बिग डेटा को संभालने में उनके लाभों को समझाइए।
----	--

b.	Describe how Zookeeper works and its role in maintaining a reliable distributed Hadoop ecosystem. Zookeeper कैसे काम करता है और वितरित Hadoop पारिस्थिति की तंत्र को विश्वसनीय बनाए रखने में इसकी भूमिका का वर्णन कीजिए।
----	---