

7E1824

Roll No. 20EMBJ7001

140411

[Total No. of Pages : 2]

7E1824

B.Tech. VII-Sem. (Main) Examination, December - 2023
Information Technology
7IT4-01 Big Data Analytics

Time : 3 Hours

Maximum Marks : 70

Instructions to Candidates:

Attempt All Ten questions from Part A, Five questions out of Seven from Part B and Three questions out of Five questions from Part C.

Schematic diagrams must be shown wherever necessary. Any data you feel missing suitably be assumed and stated clearly. Units of quantities used/calculated must be stated clearly.

Use of following supporting material is permitted during examination. (Mentioned in form No. 205).

PART - A

(Answer should be given up to 25 words only)

All questions are Compulsory.

(10×2=20)

1. What is Big Data? How does it works?
2. What are the different platforms to deal with Big Data?
3. What is the necessity of driver code?
4. List the components of a map reduce application that we can develop
5. Define the following wrappers Null writable, Object writable
6. Define the significance of comparator.
7. Define the basic syntax of a pig.
8. What are the three key design principles pig latin?
9. What is Hive Data Manipulation Language.?
10. Write at least two differences between pig and hive.

PART-B

(Analytical/Problem Solving questions)

Attempt any Five questions.

(5×4=20)

1. Discuss the problems with traditional large scale systems along with features of big data.
2. What is a Data Node? How many instances of Data Node run on a Hadoop Cluster.
3. Explain the role of combiner, record reader and partitioner within a map reduce program model of Hadoop
4. Explain the Writable class hierarchy with a neat sketch.
5. Explain with an example, how Hadoop uses scale out feature to improve the performance.
6. Explain the four data types of Pig's data model with an example.
7. How can we install the Apache Hive on the system-Explain

PART - C

(Descriptive/Analytical/Problem Solving/Design questions)

Attempt any Three questions.

(3×10=30)

1. Discuss the google file system and Hadoop Distributed File System in detail.
2. Explain about the implementation of map reduce concept with an example.
3. Explain the significance of writable interface along with writable comparable and comparators w.t.to implementing the serialization
4. Discuss about the operators supported by PIG along with PIG commands
5. Draw the architecture of Hive and explain about the various data types supported by HIVEQL with an example.