No. of Printed Pages : 4
Roll No. .....

220854B

5th Sem. Branch : Computer, Computer (For Speech and Hearing Impaired)
Subject : Big Data

Time : 3 Hrs. **M.M. : 60**

SECTION-A

Note: Multiple Choice Questions. All Questions are compulsory. (6x1=6)

Q.1 Which of the following is NOT a characteristic of Big Data?

a) Volume b) Variety
c) Velocity d) Simplicity

Q.2 What is one of the primary benefits of Big Data?

a) Increased storage costs
b) Improved decision-making through data analysis
c) Limited data accessibility
d) Slower processing speeds

Q.3 Which component of Hadoop is responsible for managing the file system namespace?

a) Data Node b) Task Tracker
c) Name Node d) Job Tracker

Q.4 What is the primary purpose of the Map Reduce frame work?

a) To create graphical user interfaces
b) To process and analyze large data sets
c) To manage network security
d) To store files in HDFS

Q.5 Which of the following best defines Big Data?

a) Data that is east to manage
b) Extremely large data sets that may be analyzed to reveal patterns and trends
c) Data stored only on local machines
d) Data that cannot be accessed

Q.6 How does Hive differ from traditional RDBMS?

a) Hive uses SQL-Like queries for data analysis.
b) Hive stores data in a relational format only
c) Hive is faster for small data sets
d) Hive does not support data partitioning

Section-B

Note: Objective/Completion type questions. All questions are compulsory. (6x1=6)

Q.7 HDFS stands for _____?

Q.8 _____ has world largest Hadoop Cluster?

Q.9 List two applications of Big Data.

Q.10 Which language is used in Hadoop?

Q.11 Data in _____bytes size is called Big Data?

Q.12 RDBMS stands for?

Section-C

Note: Short answer type Question. Attempt any eight questions out of Ten Questions. (8x4=32)

Q.13 Outline the steps of the Map Reduce word-count example, from loading data to the final output.

Q.14 What are the main characteristics of Big Data? Discuss at least three.

Q.15 Discuss how hive differs from traditional relational database management systems (RDBMS).

Q.16 What happens during the Shuffling and sorting phase in Map Reduce?

Q.17 Elaborate when Pig should be used.

Q.18 Illustrate architecture of Map Reduce.

Q.19 List different sources of Big Data.

Q.20 Describe the five daemons of Hadoop and their function alities.

Q.21 What is Pig, and how does it relate of Hadoop?

Q.22 What is HDFS and what is its role in the Hadoop ecosystem?

Section-D

Note: Long answer questions. Attempt any two question out of three Questions. (2x8=16)

Q.23 Define Big Data. What are the challenges associated with processing Big Data?

Q.24 Describe the architecture of Hadoop, including its main components and their functions.

Q25 Explain the following:

a) Map Reduce b) Hadoop

c) Pig