CBCS Scheme

USN | | | | | | | | | | |

16SCS/SCE21

Second Semester M.Tech. Degree Examination, June/July 2017
Managing Big Data

Time: 3 hrs. Max. Marks: 80

Note: ***Answer FIVE full questions, choosing one full question from each module.***

Module-1

1 a. What is Big Data and why it is important? **(08 Marks)**

b. Explain the timeline of recent technology developments with a neat diagram. **(08 Marks)**

OR

2 a. Define three V's of Big Data. **(06 Marks)**

b. List the different applications in which data can be analyzed. **(10 Marks)**

Module-2

3 a. Define impedance mismatch. **(04 Marks)**

b. Briefly explain (i) Key value and document (ii) Column family store with respect to aggregation. **(06 Marks)**

c. Mention the advantage of relational database as compared with schemaless database. **(06 Marks)**

OR

4 a. Mention the importance of graph database with a neat diagram. **(08 Marks)**

b. Explain different sharding and replication technique which can be used in NOSQL. **(08 Marks)**

Module-3

5 a. Write a java script for Mupper and reducer considering weather dataset as an example, output must retrieve maximum temperature for every year. **(10 Marks)**

b. Describe with a neat diagram Map Reduce data flow with a single reduce task. **(06 Marks)**

OR

6 a. Describe HDFS concepts:
(i) Blocks (ii) Name node and data nodes (iii) HDFS federation
(iv) HDFS high-availability. **(12 Marks)**

b. Explain Anatomy of a file read. **(04 Marks)**

Module-4

7 a. Why YARN was designed? Explain with a diagram how status up dates are propagated through map reduce in YARN. **(10 Marks)**

b. Mention atleast three output formats in Hadoop. **(06 Marks)**

OR

8 a. Briefly explain how hadoop runs a map reduce job using the classic framework. **(10 Marks)**

b. Write a segment of a program used for mock testing with respect to mapper. **(06 Marks)**

Module-5

9 a. Explain primitive and complex data types in hive. **(10 Marks)**

b. Mention the pig latin relational operators. **(06 Marks)**

OR

10 a. Write a note on Hbase data model. **(08 Marks)**

b. Mention the design differences between RDBMS and Cassandra. **(08 Marks)**

* * * * *