Name: Enrolment No:	UPES UNIVERSITY OF TOMORROW

Course: Big Data Storage   Semester : IV
Program: B.Tech. CSE (Big Data)   Time   : 03 hrs.
Course Code: CSBD2001   Max. Marks: 100

Instructions: Calculator is not allowed.

SECTION A
(5Qx4M=20Marks)

S. No.	Answer all the questions	Marks	CO
Q 1	Define the base properties of any NoSQL database elaborately	4	CO1
Q 2	Interpret the CAP theorem in brief.	4	CO2
Q 3	Illustrate the Hadoop Map-Reduce Program.	4	CO3
Q 4	Evaluate how much faster a parallel program can run using Amdahl's Law when 85% of your program can be parallelized, and 8 machines are used to run your parallel version of the program.	4	CO4
Q 5	Explain the overall architecture of GFS? Illustrate the different features of provided by GFS?	4	CO4

SECTION B
(4Qx10M= 40 Marks)

Q6	Define hard and soft links in a file system. Describe the extended partitioning.	6+4	CO1
Q7	Explain architecture of HDFS with suitable block diagram. Interpret the role of Data node and Name node in HDFS.	5+5	CO2
Q8	Illustrate Sharding and its architecture elaborately. Demonstrate Auto Sharding in MongoDB precisely. OR Illustrate the scaling of the database and its different types with suitable examples. Demonstrate the HBase Data Model in brief.	6+4 4+6	CO3

Q9	Illustrate key-value store in NoSQL. Explain the Document-Store and Column Store databases.	4+6	CO4
	SECTION-C **(2Qx20M=40 Marks)**		
Q 10	Compare RDBMS and NoSQL-based database management systems in brief. Explain INODES, Log-Based File Systems and FAT elaborately.	8+12	CO4
Q 11	Explain the various features and aspects of MongoDB. Interpret the load-Balancing architecture of MongoDB in brief. Explain High-Performance Data Access in the MongoDB environment precisely. OR Explain the following: a) HDFS block replication b) Rack awareness c) Name node high availability d) Client-server architecture in DFS	8+7+5 20	CO5