



DHARMSINH DESAI UNIVERSITY, NADIAD
FACULTY OF TECHNOLOGY
B.TECH. SEMESTER VII [CE]
SUBJECT: (CE720) BIG DATA ANALYTICS

Examination : Second Sessional
Date : 09/10/2020
Time : 10:00 a.m. TO 11:15 a.m.

Seat No : _____
Day : Friday
Max. Marks : 36

INSTRUCTIONS:

1. Figures to the right indicate maximum marks for that question.
 2. The symbols used carry their usual meanings.
 3. Assume suitable data, if required & mention them clearly.
 4. Draw neat sketches wherever necessary.
-

Q.1 Do as directed.

- (a) List out reasons why Apache Spark has become more popular. [2]
- (b) Explain in brief replication strategies of Cassandra. [2]
- (c) Explain SparkDynamicMemoryUtilization. [2]
- (d) Provide and brief about any two ways related to improved performance via storage techniques of Hive. [2]
- (e) What is SerDe and explain how does it work? [2]
- (f) Discuss various advanced data types available with Hive and Pig with their name and purpose. [2]

Q.2 Attempt *Any Two* from the following questions.

- (a) Explain using algorithm or example, Maekawa's quorum based approach to achieve Distributed Mutual Exclusion. Also explain how is quorum implemented? [6]
- (b) Explain using algorithm or example, Suzuki Kasami's token based approach to achieve Distributed Mutual Exclusion. Also explain how is token implemented? [6]
- (c) (I) Describe file formats supported by Hive with their purpose. [3]
(II) Describe Hive crud operations with syntax. [3]

Q.3 (a) Write Cassandra queries to perform following tasks: [6]

1. Create a table Product with five attributes id, name, category, price and seller, where data is partitioned on two columns name and category and sorted on id.
2. Insert a record in Product Table.
3. List out all products of category 'Groceries' and name 'Mustard'. Also make sure that the output is ordered by id.
4. Create a secondary index on seller and use seller in where clause of select query to verify your answer.
5. Check current consistency level and then set it to quorum.
6. Update seller of records having category 'Groceries' and name 'Mustard'.

- (b) Write a short note on SparkContext of Apache Spark. [6]

OR

Q.3 (a) (I) Compare and contrast Hadoop Mapreduce and Apache Spark in terms of processing. [3]

- (II) Explain different possibilities by which Spark can work together with Hadoop. [3]

- (b) Explain different types of collections in Cassandra with example. [6]