# JIS College of Engineering

(An Autonomous Institute)

Block 'A', Phase-III, Kalyani, Nadia, Pin – 741235

## FRONT PAGE

TEST – I, EVEN Semester Examination 2020-21

| | |
|---|---|
| BRANCH | Computer Science & Engineering |
| SEMESTER / YEAR | 6th semester / 3rd Year |
| REGISTRATION NUMBER | 181230110092 |
| UNIVERSITY ROLL NUMBER | 12318070 3092 |
| SUBJECT OF EXAMINATION | Distributed Database (CS 605C) |
| DATE OF EXAMINATION | 16/04/21 |
| FULL SIGNATURE OF THE CANDIDATE | Sourik Pal |

### INSTRUCTION TO THE EXAMINEES

1. Download and print this page for each examination. Fill it up and attach on the top of the answer script. Use A4 size paper. Leave back side of the front page blank.

1. Use A4 size paper to write your answers. Write answers in own hand writing.
2. Specify page number at the top of each page of the answer script.
3. Write branch name, roll number & subject name and put full signature at the bottom of each page of the answer script.
4. Do not forget to attach the front page. In absence of duly filled in front page, answer script will be treated as incomplete and will not be considered for evaluation.
5. Send the answer script along with the filled in front page to respective department.

## Group A

1)

| Centralized Database | Distributed Database |
|---|---|
| a) It is database that is stored, located as well as maintained at a single location only | a) It is a database which consists of multiple databases which are connected with each other and are spread across different physical locations. |
| b) The data access time in case of multiple users is more in a centralized database | b) The data access time in case of multiple users is less in a distributed database. |
| c) Centralized database is less costly | c) This database is very expensive. |
| d) This database provides a uniform and complete view to the user | d) Since it is spread across different locations thus it is difficult to provide a uniform view to the user |
| e) This database has more data consistency in comparison to distributed database | c) This database may have some data replications thus data consistency is less. |

Branch - CSE  Subject - Distributed Database

Roll - 12318070309 2  Signature - Sourik Pal

3) Advantages of Database Replication :-

i) To provide a consistent copy of data across all database nodes

ii) To increase the availability of data

iii) The reliability of data is increased through data replication.

Disadvantages of Database Replication :-

i) More storage space is needed as storing the replicas of same data at different sites consumes more space

ii) Data Replication becomes expensive when the replicas at all different sites need to be updated.

iii) Maintaining Data consistency at all different sites involves complex measures.

## Group-B

5) a) Data Replication is the process of storing data in more than one site or node. It is useful in improving the availability of data. It is simply coping data from a database from one server to another server so that all the users can share the same data without any inconsistency. The result is a distributed database in which users can access data relevant to their task without interfering with the work of others.

Data replication encompasses duplication of transactions on an ongoing basis, so that the replicate is in a consistently updated state and synchronized with the source. However in data replication data is available at different locations, but a particular relation has to reside at only one location.

Fragmentation should be done in a way so that the original table can be reconstructed from the fragments. This is needed so that the original table can be reconstructed from the fragments whenever required. This requirement is called "reconstructiveness".

8) Horizontal Fragmentation :-

Horizontal fragmentation groups the tuples of a table in accordance to values of one or more fields. Horizontal fragmentation should also confirm to the rule of reconstructiveness. Each horizontal fragment must have all columns of the original base table.

For example :- University Database

Student table :-

| Reg_No | Name | Course | Address | Semester | Fees | Marks |
|---|---|---|---|---|---|---|

In the student schema, if the details of all student of computer science course needs to be maintained at the school of computer Science, then the designer will horizontally fragment the database as follows :-

```
CREATE COMP_STD AS
    SELECT * FROM STUDENT
    WHERE COURSE = "Computer Science";
```

## Vertical Fragmentation

In vertical fragmentation, the fields or columns of a table are grouped into fragments. In order to maintain reconstructiveness, each fragment should contain the primary key field (s) of the table. Vertical fragmentation can be used to enforce privacy of data

The fees details are maintained in the accounts section. In this case, the designer will fragment the database as follows -

```
CREATE TABLE STD_FEES AS
    SELECT Regd_No, Fees
    FROM STUDENT;
```

6) We define mixed fragmentation as a process of simultaneously applying the horizontal and vertical fragmentation on a relation. It can be achieved in one of two ways by performing horizontal fragmentation followed by vertical fragmentation. or by performing vertical fragmentation followed by horizontal fragmentation. The need for mixed fragmentation arises in distributed databases because database users usually access subsets of data which are vertical and horizontal fragments of global relations and there is a need to process queries or transactions that would access these fragments optimally.

7) For simplifying a query :-

i) Define business requirement first
ii) SELECT fields instead of using SELECT *
iii) Avoid SELECT DISTINCT
iv) Create joins with INNER JOIN (not WHERE)
v) Use WHERE instead of HAVING to define filters

VI) [redacted] Use wildcards at the end of a phrase only

VII) [redacted] Use LIMIT to sample query results

VIII) Run your query during off-peak hours