Name : ....

Roll No. : ....

Invigilator's Signature : ....

CS/M. Tech (CSE)/SEM-2/PGCSE-202/2013

2013
ADVANCED DATABASE MANAGEMENT SYSTEM

*Time Allotted* : 3 Hours   *Full Marks* : 70

The figures in the margin indicate full marks.

Candidates are required to give their answers in their own words as far as practicable.

Answer question no. 1 and any five from the rest.

1. a) What is data cube ? Explain with example.   2

    b) Describe the types of skew handling in parallel DBMS. 2

    c) Describe the types of data marts.   2

    d) Describe the basic mechanisms of independent parallelism, pipelined parallelism and partitioned parallelism for parallel join query execution.   1 + 1 + 1

    e) Describe the shared memory, shared disk, and shared nothing architectures of parallel database.   1 + 1 + 1

    f) Write the sorted merge join algorithm and hash split join algorithm for parallel query execution.   2 + 2

    g) Describe the steps for the local transaction and global transaction of distributed DBMS architecture.   2 + 2

2. a) Describe the range, round-robin and hash partitioning algorithms and comparison of these partitioning algorithms for parallel database.

30230 (M.Tech)   [Turn over
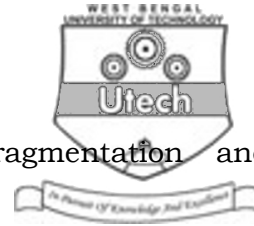
b) Consider the following global, fragmentation and allocation schema :

Global Schema :

STUDENT (STUD-ID, DEPT, NAME)

Fragmentation Schema :

$$STUDENT1 = SL_{DEPT="EE"} STUDENT$$

$$STUDENT2 = SL_{DEPT="CS"} STUDENT$$

Assume that EE and CS are only possible values for DEPT.

Allocation Schema :

STUDENT1 at sites 1, 2

STUDENT2 at sites 3, 4

i) Write a query that accepts the STUD-ID from terminal and outputs the name and the department at levels 1, 2 and 3 of transparency.

ii) Write a query that moves the student having number 232 from department "EE" to department "CS" at levels 1, 2 and 3 of transparency. 6 + 4

3. a) What is fragmentation transparency ? Explain with examples of types of fragmentation transparency for distributed DBMS.

b) Consider the following schemas :

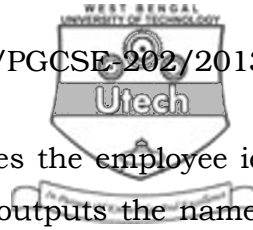Global schema : Employee (emp-id, name, dept, design, salary)

Fragment schema :

$$EMP1 = \sigma_{dept = "DEVELOPMENT"} (Employee)$$

$$EMP2 = \sigma_{dept = "PRODUCTION"} (Employee)$$

Allocation schema : EMP1 at sites 1, 2 and EMP2 at sites 3, 4.

Assume that "DEVELOPMENT" and "PRODUCTION" are only possible values for the dept.

    i)   Write an application that requires the employee id (emp-id) from the terminal and outputs the name, design, salary and department at levels fragment, location and local mapping transparency.

    ii)   Write an application that moves the employee having emp-id 56 from the department "DEVELOPMENT" to the department "PRODUCTION" at different levels of data distribution transparency.

    iii)   Write an application that moves an employee whose emp-id and dept is given at the terminal to the other department at location transparency.

$$4 + (3 \times 2)$$

4.   a)   Describe the basic steps for distributed query processing.

    b)   What is a Global Query optimization ? Compare algorithm for global query optimization. Why global query optimization is difficult in distributed DBMS ?
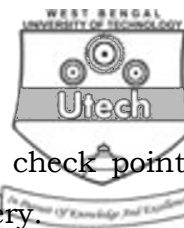
    c)   Consider the following schemas :

EMP (ENO, ENAME, TITLE)

ASG (ENO, PNO, RESP, DUR)

PROJ (PNO, PNAME, BUDGET, LOC)

Further consider the following query : SELECT ENAME, PNAME, FROM EMP, ASG, PROJ WHERE ENAP.ENO = ASG.ENO AND ASG.PNO=PROJ.PNO AND (TITLE = "ELECT"> ENG OR ASG.PNO<"P_3"). Draw the generic query tree. Transform the generic query tree to an optimized reduced query tree.   2 + 4 + 4

5. a) What is check pointing ? Explain how check pointing reduces the overhead of log-based recovery.

   b) Compare blocking and non-blocking protocols. Describe non-blocking two-phase commit protocol. What are the demerits of this protocol ? 4 + (2 + 4)

6. a) How you perform deadlock management in a distributed DBMS ? Discuss.

   b) Differentiate between preemptive and non-preemptive methods for distributed deadlock prevention.

   c) Explain the algorithm for distributed deadlock avoidance.

   d) Describe quorum-based protocol for distributed concurrency control. 2 + 2 + 3 + 3

7. a) Describe the Data Warehouse architecture.

   b) What are the advantages of data warehouse ?

   c) What is an OLAP ? What are the differences between OLAP and OLTP ?

   d) What are the differences between OLTP and data warehouse ? 3 + 2 + 3 + 2

8. a) What are the basic steps of data mining for knowledge discovery ?

   b) What are the different processes in knowledge discovery in databases ?

   c) What is an On Line Analytical Mining (OLAM) ? Describe the OLAM architecture.

   d) What is difference between data mining and OLAP ?

   2 + 2 + 4 + 2