

**Eighth Semester B. E. (Computer Science and Engineering)  
Examination**

**Elective – III**

**DISTRIBUTED AND PARALLEL DATABASES**

Time : 3 Hours ]

[Max. Marks : 60

**Instructions to Candidates :—**

- (1) All questions carry marks as indicated.
- (2) Assume suitable data and illustrate answers with neat sketches wherever necessary.

1. (a) Differentiate on the basis of various features between centralized and distributed databases. 5 (CO 1)  
(b) Sketch the following parallel database architectures and discuss their advantages and disadvantages :—
  - (1) Shared Memory.
  - (2) Shared Disk. 5 (CO 1)
2. (a) Consider the following relations :—  
BOOKS(Book#, Primary\_author, Topic, Total\_stock, \$ price)  
BOOKSTORE(Store#, City, State, Zip, Inventory\_value)  
STOCK(Store#, Book#, Qty).  
Total\_stock is the total number of books in stock and Inventory\_value is the total inventory value for the store in dollars. Design a global schema for above database.
  - (i) Give an example of two simple predicates that would be meaningful for the BOOKSTORE relation for horizontal partitioning.
  - (ii) How would a derived horizontal partitioning of STOCK be defined based on the partitioning of BOOKSTORE ?
  - (iii) Show predicates by which BOOKS may be horizontally partitioned by topic. 5 (CO 2)

- (b) Describe the components of reference architecture for a distributed database. 5 (CO 2)

3. Solve any **Two** :—

- (a) Discuss the concurrency control mechanism based on Locking in Distributed databases. 5 (CO 2)
- (b) Explain the conservative timestamp mechanism. 5 (CO 2)
- (c) Consider a data item  $x$ . Let  $RTM(x) = 35$  and  $WTM(x) = 40$ . Let the pair  $\{R_i(x), TS\}$  ( $\{W_i(x), TS\}$ ) denote a read(write) request of transaction  $T_i$  on the item  $x$  with timestamp  $TS$ . Determine the behavior of the basic timestamp method with the following sequence of requests :
- $\{R_1(x), 37\}, \{R_2(x), 22\}, \{W_3(x), 15\}$   
 $\{W_4(x), 26\}, \{R_5(x), 31\}, \{W_6(x), 42\}$  5 (CO 2)

4. Solve any **Two** :—

- (a) How reduction of relations are possible using a semi join ? Explain with example. 5 (CO 3)
- (b) While doing equivalence transformation of queries, list the various criteria being used. Explain any three of them with an example. 5 (CO 3)
- (c) Assume that the fragmentation of a global relation  $R(A, B, C)$  is described using the following predicates :

$P_1 : (1 \leq A \leq 4)$

$P_2 : (5 \leq A \leq 7) \text{ and } (1 \leq B \leq 5)$

$P_3 : (5 \leq A \leq 7) \text{ and } (6 \leq B \leq 10)$

$P_4 : (8 \leq A \leq 10) \text{ and } (1 \leq C \leq 5)$

$P_5 : (8 \leq A \leq 10) \text{ and } (6 \leq C \leq 10)$

Introduce the CUT operation in the following parameteric queries :

(i)  $SL_{A=\$X} \text{ and } B<\$Y R$

(ii)  $SL_{C=\$X} \text{ and } B=\$Y R$

(iii)  $SL_{A>\$X} \text{ OR } C=\$Y R$

5 (CO 3)

5. (a) Discuss the concept of a distributed transaction, LTM and DTM with respect to distributed transaction recovery. 5 (CO 2)
- (b) Voting is a technique for managing concurrent data access. Explain the working of voting based protocol. 5 (CO 2)
6. (a) List the various algorithms in centralized approach for data mining. Explain any one algorithm in detail. 5 (CO 4)
- (b) Describe the following approaches for distributed and parallel data mining :—
- (i) Count distribution.
  - (ii) Data distribution.
  - (iii) Candidate distribution.
  - (iv) eclat. 5 (CO 4)