Total No. of Questions : 5]

SEAT No. :

P1786

[Total No. of Pages : 5

[5233]-13

M.Sc

COMPUTER SCIENCE

CS - 103 : Distributed Databases Concepts

(2008 Pattern) (Semester - I)

Time : 3 Hours]

[Max. Marks : 80

Instructions to the candidates:

1) *All questions are compulsory.*
2) *All questions carry equal marks.*
3) *Figures to the right indicate full marks.*
4) *Assume suitable data if necessary.*

Q1) Attempt any eight from the following. **[8×2=16]**

a) Give any two complicating factors in design of DDBMS.

b) Explain what is unilingual architecture.

c) Write short note on distributed catlog management.

d) What are the query processing objective.

e) Explain the term.

   i) Type incorrect query.

   ii) Semantically incorrect query.

f) Define

   i) Linear join tree

   ii) Bushy join tree

g) List any four types of failures that can occure in distributed databases.

h) What is nested transaction?

i) Discuss the classification criteria of concurrency control approaches.

j) What are the layers of query processing?

P.T.O.

Q2) Attempt any four from the following. **[4×4=16]**

    a)   Explain steps of query processor by specifying input and output of each step.

    b)   Providing transparency is necessary in DDBMS. Comment.

    c)   Autonomy refers to distribution of control, not data. Comment.

    d)   Compare and contrast between following two architectures of DDBMS.

        i)   Client - server

        ii)   Peer - to - peer

    e)   Replication of data in distributed databases reduces reliability, justify true or false.

Q3) Attempt any four from the following. **[4×4=16]**

    a)   Consider the following database :

        PROJ (pno, pname, budget, location) and

        ASG (pno, eno, duration, responsibility)

    Consider, set of simple predicates

    P = {budget $\leq$ 100000,

        budget >100000}

    Perform horizontal fragmentation of PROJ based on set *p*, Using this fragmentation of PROJ,

    Further perform derived horizontal fragmentation of ASG.

    b)   Given database

        EMP (<u>eno</u>, ename, title) and

        ASG (<u>eno</u>, <u>pno</u>, dur, resp)

    Supose EMP relation is horizontally Fragmented as

        $EMP_1 = 6$ eno $\leq$ "e_3" (EMP)

        $EMP_2 = 6$ eno > "e_3" (EMP)

Supose ASG relation is horizontally fragmented as

$ASG_1 = 6$ eno $\leq$ "e_3" (ASG)

$ASG_2 = 6$ "e_3" < eno $\leq$ "e_5" (ASG)

$ASG_3 = 6$ eno > "e_5" (ASG)

Transfer following query to reduced query on fragments and determine whether it is better than generic query

SELECT ename, title, resp, pno
FROM ASG, EMP

WHERE ASG. eno = EMP. eno and pno = " $p\,l$ " ;

c) Consider following query :

SELECT patient - name

FROM patient P. disease D. patient - disease PD

WHERE P. sno = PD. sno

and PD. dno = D. dno

and dname = "DENGUE"

and P. age < 55

Optimize above query using centralized INGRES query optimization algorithm.

d) Consider query that refers to join of relation.

PROJ (pno, pname, bud, 10c) &

ASG (pno, eno, dur, resp) on attribute

pno.

Assume that PROJ & ASG reside at two different sites, and index is on pno.

Consider size (PROJ) = 200 and size (ASG) = 400. Write down all possible execution strategy along with cost incurred. If distributed system R algorithm (R*) is implemented. Which is best strategy amongst all and why?

e) Consider the following distributed wait - for - Graph
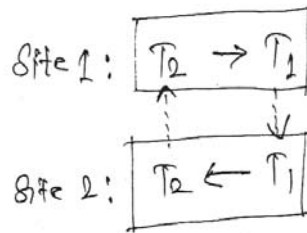
Detect Deadlock using distributed deadlock detection algorithm.

Q4) Attempt any four from the following. **[4×4=16]**

a) Consider the following relation

EMP (eno, ename, addr, age, dno)

DEPT (dno, dname, bud)

EMP relation is partitioned horizontally as

EMP_1 = 6 age < 25 (EMP)

EMP_2 = 6_{25} < age < 50 (EMP)

EMP_3 = 6 age > = 50 (EMP)

DEPT relation is partitioned horizontally as

$DEPT_1$ = 6 bud < 3,00,000 (DEPT)

$DEPT_2$ = 6 bud > = 3,00,000 (DEPT)

i) Draw join graph of EMP ⋈ DEPT

ii) Is the graph simple or partitioned?

b) There are two relations in DDBMS.

EMP (eno, enm, age, sal, desg, dno)

DEPT (dno, dnm, bud)
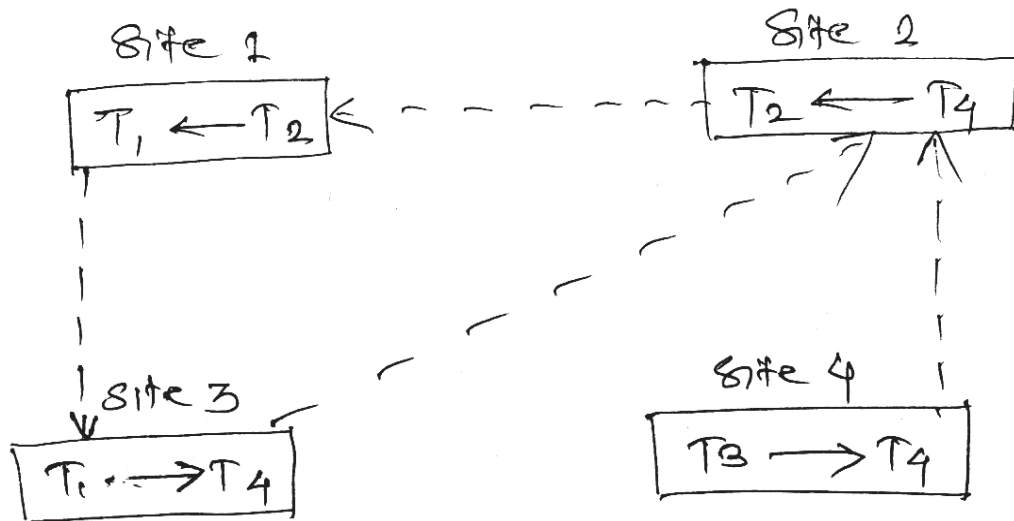
Size (EMP) = 10,000 taples

Size (DEPT) = 500 taples

EMP is stored at site S_1

DEPT is stored at site S_2

Query executed at site S_3

"Find out names of managers of those departments which have budget i.e bud > 7,00,000.

c) Consider following DWFG.

Check if deadlock exist in system. If so find out the sites involved in deadlock.

d) GNe schedule of two complete transaction which is not allowed by strict 2 PL, but is allowed by basic 2PL.

e) In 3PC, when new coordinator is selected due to failure of original coordinator how it handles the termination protocol?

Q5) Attempt any four from the following. **[4×4=16]**

a) Write note on : Work flow.

b) Explain each rule of fragmentation correctness in detail.

c) Explain optimistic concurrency control protocol.

d) Explain centralized 2PC protocol in distributed environment.

e) Write a short note on "False deadlocks".

✡ ✡ ✡