CS5225 Parallel & Distributed Databases

Assignment 2

(Due: 22 Oct 2007)

1. Consider the following schedules:

S_1=$W_2(x)$, $W_1(x)$, $R_3(x)$, $R_1(x)$, $W_2(y)$, $R_3(y)$, $R_3(z)$, $R_2(x)$
S_2=$R_3(z)$, $R_3(y)$, $W_2(y)$, $R_2(z)$, $W_1(x)$, $R_3(x)$, $W_2(x)$, $R_1(x)$
S_3=$R_3(z)$, $W_2(x)$, $W_2(y)$, $R_1(x)$, $R_3(x)$, $R_2(z)$, $R_3(y)$, $W_1(x)$
S_4=$R_2(z)$, $W_2(x)$, $R_3(z)$, $W_1(x)$, $W_2(y)$, $R_1(x)$, $R_3(x)$, $R_3(y)$

a) Which of the above schedules are conflict equivalent?
b) Which of the above schedules are serializable?
c) Suppose the each schedule is located at an individual site. Each site has replicas of items x, y and z. Consider those schedules that are local serializable. Are they globally serializable as well?

2. Assume that the data are distributed across 8 sites A to H. Item a is stored at site A, item b is stored at site B and so on. $R_i(a)$ and $W_i(a)$ denote read and write lock request by transaction *i* on data item a. Consider the following sequence of operations:

$W_1(c)$, $R_7(b)$, $R_4(a)$, $W_3(d)$, $R_8(g)$, $R_2(f)$, $W_3(g)$, $W_6(e)$, $R_5(f)$, $W_7(f)$, $R_6(c)$, $W_6(h)$, $W_8(a)$, $W_2(b)$, $R_4(d)$, $W_1(e)$

a) Draw the wait-for graph (WFG) for the above operations. Identify deadlocks if there is any.
b) Suppose the 8 sites are organized hierarchically in a binary tree for deadlock detection. Find the best organization of the sites for this scenario and explain the reason why it is the best.

3. Consider a distributed system with N participants and 1 coordinator.
a) Evaluate the three communication topologies for 2PC: centralized, linear and distributed 2PC in terms of the number of messages and the number of rounds required for these 2PC topologies.
b) For centralized 3PC, how many messages and rounds are needed?
c) In 3PC, is it possible for a node to be in PRECOMIT state and eventually be aborted? If not, explain the reason, otherwise describe such situation.
* Assume that there is no node failure and ignore the DONE messages.

4. In a shared nothing database system with 4 homogeneous nodes, relation R(X,Y) is partitioned on the attribute X across node 1 and node 2, while relation S(Y,Z) is partitioned across node 3 and node 4 on attribute Y. We want to perform equal-join between R and S on these 4 nodes. The query optimizer has the following statistical information:

Node	Table	Number of rows	Min Y value	Max Y value
1	R	16K	1	4000
2	R	24K	4001	16000
3	S	24K	1	12000
4	S	16K	12001	16000

Assume that data in each bucket is uniformly distributed.

a) If communication cost is the only concern, suggest a scheme to reallocate R and S. Calculate total the communication cost. Communication cost is measured in the number of tuples transmitted.

b) Suppose load-balance is the major concern. Suggest a scheme to reallocate R and S that minimize the communication cost while achieving the best balance between nodes. The workload is defined as the number of join results estimated.

c) After redistribution, partitions of R and S are hash-joined locally in each node. Each node has a hash table with 1000 entries. If the workload is defined as the number of comparisons during the hash-join, calculate the workload (total number of comparisons) in the above two schemes. Suppose the hash functions used are completely random.