

CMSC 724, Spring 2015: Homework 4

Due Friday April 3, 2015, 11:59pm.

The homework is to be done by yourself. You are free to make reasonable assumptions about the problem if something is not fully specified – make sure to state any assumptions. Try to be concise in your answers. You can turn in a hard copy if you'd like before the deadline (slide under my office door if I am not there). Total weight = 5%.

1. (BIRCH) Briefly explain how BIRCH handles the case when it runs out of memory because of choosing a very small threshold value.
2. (SPRINT) Briefly discuss how the *parallel* version of the SPRINT algorithm can be mapped to MapReduce framework (Section 3).
3. Consider the following history with three transactions that is permitted under Snapshot Isolation:

$$R_2(X_0, 0), R_2(Y_0, 0), R_1(Y_0, 0), W_1(Y_1, 20), C_1, R_3(X_0, 0), R_3(Y_1, 20), C_3, W_2(X_2, -11), C_2$$

T_1 simply deposits \$20 into one account (Y), T_2 withdraws \$10 from the account X and is forced to pay a \$1 overdraft fee (because there is not enough money in the two accounts – if there were more than 10 dollars in the two accounts combined, the overdraft fee is not charged), whereas T_3 simply reads and prints the balances. Explain concisely why this schedule is not serializable. Why would it not be allowed by 2-phase locking? Would OCC allow this schedule?

4. (Optimistic Concurrency Control) Briefly in your words explain the optimization (what is the optimization, and why is it desired) discussed at the end of Section 4 that is centered around using "mid tn ".
5. Briefly explain under what circumstance the algorithm by Lehman-Yao will hold 3 locks simultaneously (this is discussed on Page 662).