Assignment 3

Joining Techniques

Submission Date: Tue 17-Oct-2017 (Start of Class)

Instructions:

- Read the assignment carefully and thoroughly.
- If you have any confusion in understanding the statement then make suitable assumptions
- Mention your assumptions before solving the question.

Question 1. Discuss in which particular case/cases you will prefer

- Nested Loop Join (NLJ)
- Sort Merge Join (SMJ)
- Hash Join (HJ)
  You need to justify, why this technique is optimal for this case/cases.

Instructions for Question 2: You will have to take assumptions for block size (B), record size (R), number of records (r), available memory (K), indexing column, index type (clustered/non-clustered), and all remaining factors required in calculations.

Question 2. Consider two tables A and B which have to be joined. Calculate the cost of joining the two tables on their common attribute. You will have to consider the following cases and have to calculate cost for all these cases:

1. When the table sizes are almost similar. Let's say 1 row of table A joins with only 1 row of table B.

- When RAM size is not sufficient.
- When RAM size sufficient for both tables.

2. When the sizes of both tables are significantly different. Let's say 1 row of table A joins with 10 rows of table B.

- When RAM size is not sufficient.
- When RAM size sufficient for both tables.

Your solutions should include costs of I/Os for **Naïve nested loop join (or Block NLJ/Index NLJ/clustered index NLJ), Sort merge join (or merge join), and hash join**.