

INF 551 – Spring 2018

Quiz 9: Query execution (10 points), 15 minutes

Suppose $M = 101$ pages, $B(R) = 2,000$ blocks, $B(S) = 5,000$ blocks. Consider $R \bowtie S$.

1. [5 points] What is the cost (i.e., the number of block I/O's) of performing the join using the **sort-merge** join algorithm? Assume that we use 100 pages in sorting. Explain your answer (by showing the cost at each step of the algorithm).

- a. Sort R into sorted runs of length 100 and send them back to disk: Cost = $2 B(R) = 4,000$
- b. Sort S into sorted runs of length 100 and send them back to disk: Cost = $2 B(S) = 10,000$
- c. Join all sorted runs: Cost = $B(R) + B(S) = 7,000$

Total cost: 21,000 Block I/O's.

2. [5 points] What is the cost of performing the join using the **partitioned-hash** join algorithm? Explain your answer as well.

- a. Hash R into $M - 1 = 100$ buckets and send them back to disk: Cost = $2 B(R) = 4,000$
- b. Hash S into $M - 1 = 100$ buckets and send them back to disk: Cost = $2 B(S) = 10,000$
- c. Read each bucket of smaller relation into memory. For each bucket, read the same bucket of the larger relation block by block and join the matching tuples. Cost: $B(R) + B(S) = 7,000$

Total cost: 21,000 Block I/O's.