Quiz 6 Answer, Fall 2019
CSCI 4380 DB Sys

Question 1 (12 points). You are given a relation `T` where `PAGES(T)=80,000`.

What is the cost of sorting `T` if $M = 200$?

Answer.

Step 1: Cost = 160K (read and write), Creates 80,000/200 = 400 sorted groups
Step 2: Cost = 160K (read and write), Reduces 400 sorted groups to 2 (400/200)
Step 2: Cost = 80K (read and output)

Total cost = 400K pages

Question 2 (7 points). You are given the following information for relations `R` and `S`:

`TUPLES(R)=100,000, PAGES(R)=2,000, TUPLES(S)=10,000,000, PAGES(S)=50,000`

What is the cost of joining this relation if $M = 401$?

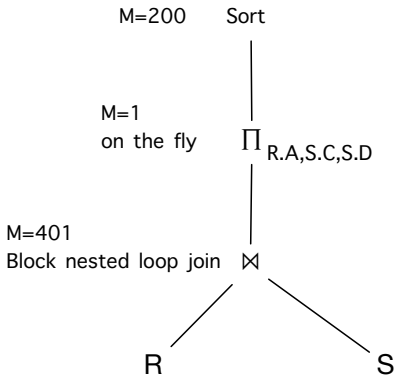
Answer. Using block-nested loop join, reading 400 pages of R and 1 page of S.

Read R once, Read S: 2,000/400 = 5 tuples

Total cost = 2,000 + 50,000*5 = 252K pages

Question 3 (6 points). Suppose you are given the query plan in the figure below. Note that: $T = \Pi_{R.A,S.C,S.D}(R \bowtie S)$ (same `T` from Question 1 and `PAGES(T)=80,000` as in Question 1).

Based on this, compute the overall cost of this query and show your work.



Answer. Basically, the sort reads the relation from the operation below and projection has no additional cost, saving the first 80K pages of cost for sort.

Total cost: 252K (join) + 400K (sort) - 80K (no initial read due to pipelining) = 572K pages