CSCI3170 Short Assignment #4

(Deadline: Dec 4 23:59)

Name: Pass / Fail

Student ID:

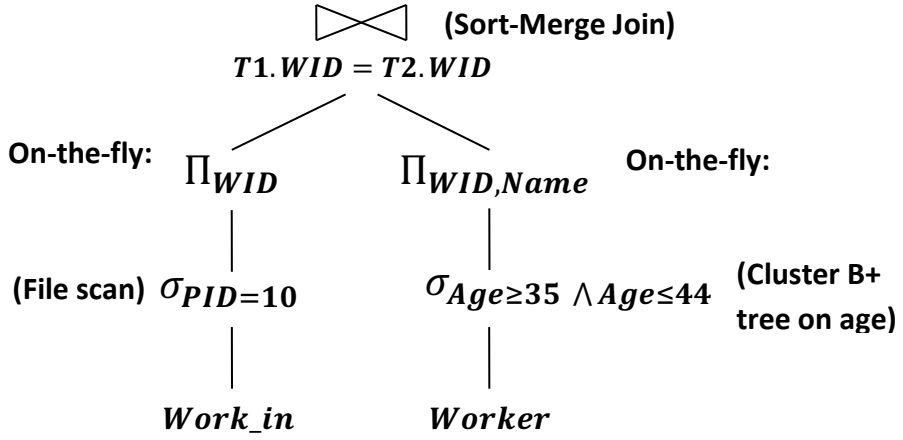
Consider the following three relations for the database in a company:

Worker (*WID*: integer, *Name*: string, *age:* integer, *rating*: integer)
Project (*PID*: integer, *project_name*: string, *budget*: real)
Work_in (*PID*: integer, *WID*: integer, *since:* date)

Assumptions:

- Assume $21 \leq age \leq 60$
- Assume $1 \leq \text{rating} \leq 10$
- The distribution of *age* in **Worker** is uniform.
- The distribution of *rating* in **Worker** is uniform.
- **Worker:** 50 tuples per page, 1000 pages
- **Project**: 60 tuples per page, 2000 pages
- **Work_in** : 40 tuples per page, 5000 pages
- The buffer size is 20 pages.

Consider the following execution plan. Assume the sizes of T1 and T2 are 20 and 50 pages respectively.



i) Calculate the number of page accesses for scanning the table **Work_in** and writing the matching tuples to **T1**.

ii) Assume the height of the B+ tree on age is 3 (i.e. number of levels = 4). Calculate the number of page accesses for selecting the tuples from **Worker** and writing the matching tuples to **T2**.

iii) Calculate the number of page accesses for sorting **T1** and **T2**, and hence calculate the number of page accesses for the Sort-Merge Join of **T1** and **T2**.
[Hint: When M > B, the formula for the cost of sorting is $\mathbf{2*M*(\lceil \log_{B-1} M/B \rceil + 1)}$]