

CSCI3170 Short Assignment #4 (Solution)

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)

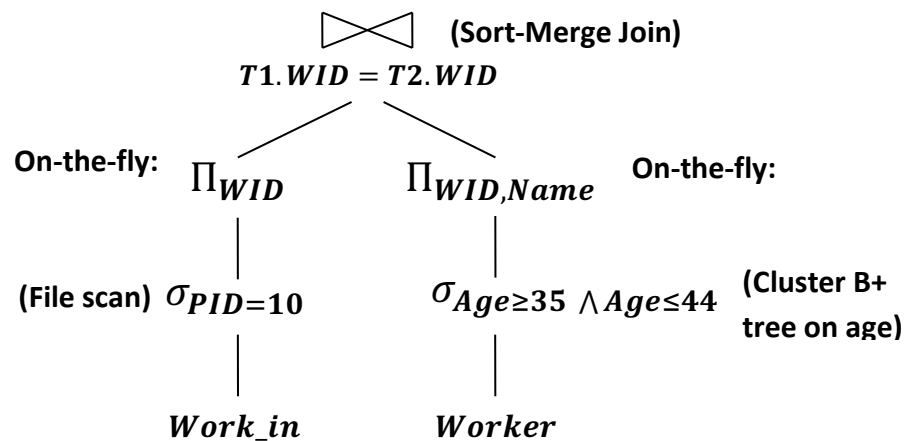
Project (PID: integer, project_name: string, budget: real)

Work_in (PID: integer, WID: integer, since: date)

Assumptions:

- Assume $21 \leq \text{age} \leq 60$
- 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**.

$$\begin{aligned} & 5000 + 20 \\ & = 5020 \end{aligned}$$

- 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**.

$$\begin{aligned} & 4 + \frac{10}{40} \times 1000 + 50 \\ & = 304 \end{aligned}$$

- 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 $2 \cdot M \cdot (\lceil \log_{B-1} M/B \rceil + 1)$]

$$\text{Sorting T1: } 2 \times 20 \times \left(\left\lceil \log_{19} \left(\frac{20}{20} \right) \right\rceil + 1 \right) = 40$$

$$\text{Sorting T2: } 2 \times 50 \times \left(\left\lceil \log_{19} \left(\frac{50}{20} \right) \right\rceil + 1 \right) = 2 \times 50 \times (1 + 1) = 200$$

$$\text{Merge Join: } 20 + 50 = 70$$

$$\text{Total: } 40 + 200 + 70 = 310$$

