

Homework 10

Problem 1:

50 pts

A relational database table has currently a size of 3.4 GB (3.4×2^{30} B.) It is stored in Flash memory with 4KB pages as storage. The average size of a record in the table is found to be 160B. We are processing a combination of a projection and a selection.

- (1) How many records are in the table?
- (2) Calculate the number of pages accessed during a table scan.
- (3) Calculate the average number of records per page. We assume that we do not store partial records in a page.
- (4) Assume that the table has an index on the attribute on which we select. The index is organized as a B+-tree. However, this is a secondary index and records in the same leaf can be in many different pages. Assume that the selection chooses 1% of all records. How many pages will an indexed search access? (You cannot answer this definitely based on the information given, but you can use an approximation.)

Problem 2:

20 pts

- (1) Express the SQL query

```
SELECT B, C
FROM R, S
WHERE A=5 AND D=10 AND R.C = S.C
```

which refers to tables R(A,B,C) and S(C,D), **literally** as an expression tree (with a cross-product).

Hint: https://link.springer.com/referenceworkentry/10.1007/978-1-4899-7993-3_293-2

- (2) Now push projections and selects as far down as possible and give the expression tree.

30 pts

Problem 3:

- (1) In the classicmodels database from mysqltutorial.org, identify the two largest tables.
- (2) When you do a horizontal fragmentation of the largest table, what other tables should you fragment horizontally as well?