

Basi di Dati, Assignment 3, 2019/20

Due: Dec 15, 4:00pm

Database

Professor (id:int, name:char(50), address:char(50), age:int, height:float)
Course (cid:char(25), title:char(50), area:char(30), instructor:int)

- instructor is a FK to Professor
- No attribute can be null.

Task

Create a java program, or a python if you prefer, that connects to your database and performs the following operations in order:

1. Drops the two tables from the database if they already exist.
2. Creates the two tables as described above.
3. Generates 1 million (random) tuples, so that each tuple has a different value for the height attribute, and insert them into the table `Professor`. Make sure that the last inserted tuple, and only that, has the value 185 for the height attribute.
4. Generates 1 more million (random) tuples and inserts them in the table `Course`.
5. Retrieves from the database and prints to `stderr` all the `id` of the 1 million professors.
6. Updates all tuples that have value 185 as `height` and makes them to have a `height` equal to 200 – (your query should work even if there are many tuples that have value 185 in the attribute `height`).
7. Selects from the table `Professor` and prints to the `stderr` the `id` and the `address` of the professors with `height` 200.
8. Creates a B+tree index on the attribute `height`.
9. Retrieves from the database and prints to the `stderr` the `id` of the 1 million professors.
10. Updates all the tuples that have value 200 as `height` and makes them to have a `height` equal to 210 -- (your query should work even if there are many tuples that have value 200 in the attribute `height`).
11. Selects from the table `Professor` and prints to the `stderr` the `id` and the `address` of the professor with `height` 210.

For each of the above operations you need to report (print to the `stdout`) the time it took to execute it. To do so you may keep in a variable the time before starting the execution (in nanoseconds), then get the system time after the execution has been completed and the difference in nanoseconds is the approximate time it took for the step to be executed. Your standard output stream should be of the form:

```
Step 1 needs 10 ns
Step 2 needs 27 ns
Step 3 needs 77 ns
...
```

Notes & Delivery

- You need to deliver ONE file only (java or python) that is named as A3_XXX.java (or .py) where the XXX is your matricola.
- The delivery is done via the Google Classroom. The instructions on how to do the delivery of an assignment to google classroom are the same as in assignment one, so you are already familiar with it.
- You may assume JDBC to be available in the classpath (java) and pycopg2 to be already installed (python).