Prof. Dr. Prof. h.c. Peter Peinl

Verteilte Datenbanken
Distributed Databases

Assignment 3

This assignment consists of a (partial) implementation of two SQL queries using the Java language and JDBC. In a real database environment those and other queries are done by the SQL query processor (e.g. in Oracle, DB2, MS SQL server). This assignment prepares the mini project to be carried out in groups of 3 or 4 students as the final task of the course.

The queries will be run on a set of simple tables have been created and loaded with synthetic data. The tables are very similar in structure and content to those created by your programs in Assignment 2. However, the tables used in this assignment are the same for every student. Tables named R1K, R10K, R100K and R1000K have already been created and loaded with data. They are stored on user id PROJA14. Furthermore, access is protected by password

However, read access has been granted to everybody (PUBLIC). You may use the tables in the SELECT statement by their system wide names, i.e. PROJA14.R1K, etc.

The overall task of the assignment is to write a java program that executes the following two SELECT statements on the tables (Rxk stands for R1K, R10K, …,, respectively) in two/three different ways:

1. SELECT COUNT (DISTINCT FIBO) FROM RxK
2. SELECT COUNT (*) FROM RxK R1, RxK R2
   WHERE R1.PK = 2 * R2.PK

The results of the two statements shall be calculated in 3 different ways.

1. Full database mode: send the entire statement via JDBC to the database.
2. No database mode: send the statement SELECT * FROM table to the database and calculate the result in your Java program.
3. Partial database mode: send the following SELECT statement to the database and calculate the result in your Java program.
   3.1. SELECT FIBO FROM table
   3.2. SELECT PK FROM table

Measure the elapsed time for each (System…..) execution in the different modes.

Compare the results to check for correctness (equality) and print them. Also print the elapsed time for each run.