

# CSCI4370 Database Management

Spring 2013

## Project 5: Performance Tuning of SQL Queries

**Due:** April 15, 8 am.

In this project, we want to test the two DBMS to see which one is faster using the queries given below.

### Queries:

1. List the name of the student with id equal to v1 (id).
2. List the names of students with id in the range of v2 (id) to v3 (inclusive).
3. List the names of students who have taken course v4 (crsCode).
4. List the names of students who have taken a course taught by professor v5 (name).
5. List the names of students who have taken a course from department v6 (deptId), but not v7.
6. List the names of students who have taken all courses offered by department v8 (deptId).

In these queries, v1, v2, ... stands for value1, value2, ...etc.

To be able to test the two database systems, which are MySQL and PostgreSQL, you need to analyze query plans generated by the DBMS and tune these queries to be able to run them faster. For tuning these queries, links to tuning guides for each DBMS are given below;

MySQL: <http://dev.mysql.com/doc/refman/5.6/en/optimization.html>

PostgreSQL: <http://www.postgresql.org/docs/9.1/interactive/internals.html>

You will need to run the queries, look at the query plans, and use some of hints given in the tuning guides of each DBMS above (for example, removal of redundant parenthesis in the query, changing order of operators, etc.). Also you can add indexes to try to speed it up. Then you need to write *a report about each query*, which is fairly short that indicates what you did to speed it up. For doing this, you need to obtain initial timing results and what the timing result is after you tuned it.

Also you need to create a database, which you can run those queries on. You'll create a sample database used in the textbook, and you can find the details of this database, which is used throughout the book (database on students, courses etc.). However, if

you just create a database with small number of tuples it is going to take 0 times, and you will not be able to make any comparison. Hence you need to insert enough tuples (e.g., 50,000 to 100,000) to get the proper timing result and be able to make a comparison in between. To do so, you need to use Tuple Generator that you have used in Project 4. You will retarget the Tuple Generator to be able to quickly populate your database.

You need to turn in before and after .sql files and query plans for six queries (given in English, see above). Do this for both the MySQL and PostgreSQL DBMSs. Present tuning experience and performance results including plots in class.

**Programming language:** Java 7 is required for the project.

**What to submit:** Please submit

- Report
- Before and after .sql files and query plans for six queries for both the MySQL and PostgreSQL
- A readme file
- Performance diagrams in Excel or PDF

The readme file should include: your names, how the workload is split (who did which part), what the zip file contains. Please pack all your files in a zip package with the file name: "project5" + last names of group members. For example: project5\_chen\_kim\_luo\_lee.zip

**How to submit:** Mail your ".zip" file to the TA (Ugur Kursuncu, [kursuncu@uga.edu](mailto:kursuncu@uga.edu))

An electronic copy of this project description and Java source code templates to be used can be found at course web page.