

Tecnologias de Base de Dados Assignment #1 | Query Optimisation

Delivery Week: 7

ASSIGNMENT GOALS

Analyse several SQL execution plans in a test database. Assess the impact of indexes and statistics and the use of different strategies for query organisation.

WORKGROUP

The work should be done in groups of one or two.

SITUATION

Take the following database model fragment about the distribution of teaching services in a faculty (Case Teaching Service). There are courses (table XUCS), described by a code (codigo), a designation (designacao), an acronym (sigla_uc) and a program (curso). Courses have occurrences in several years. Each occurrence is recorded by a row in the table XOCORRENCIAS, with information on the course code (codigo), academic year (ano_letivo), period of classes (periodo, that may be A-annual, 1S- first semester, 1T- first trimester, etc.), number of enrolled students (inscritos), students with distributed assessment (com_frequencia), number of approved (aprovados), course goals (objetivos) and content (conteudo), and department in charge (departamento).

Each occurrence may have one or more class types (T-theoretic, P-practical, L-laboratory, TP-theoretic/practical, OT- tutorial guidance). Each class type for an occurrence is recorded in table XTIPOSAULA with the number of similar classes (turnos), the number of weekly hours for each class (horas_turno), and, in some cases, the number of weekly classes (n_aulas).

Table XDSD records the teaching service distribution in each semester for each professor. More specifically, it records, for each class type of occurrence, how many weekly hours are assigned to that professor. If a professor is teaching, in a single class, more than one course at the same time, for example, from different programs, the weight of that course, in the professor's perspective, may be less than 1 and recorded in the attribute fator. Otherwise, the attribute 'fator' will be 1. From the program perspective, the attribute 'fator' is ignored. The attribute 'ordem' enables listing the set of professors of a specific course occurrence in a specific order.

The professors are recorded in the table XDOCENTES with a number (nr), a name (nome), an acronym (sigla), a category code (categoria), a given name (proprio), a family name (apelido), and a status (estado: A-ativo, NA-não ativo, R-reformado).

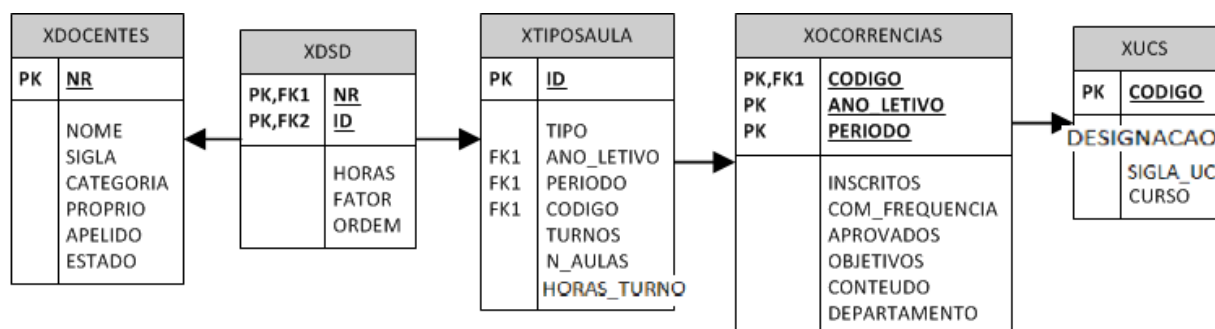


Figure 1 – Relational model for the case Teaching Service.

DATA

The tables of this database are available in the user GTD10 in the Oracle server (connection: BD, user: <user>, password: <pass>, host: oraalu.fe.up.pt, port: 1521, SID: ALU). The tables should be copied three times, with prefixes “x”, “y” and “z” (xdocentes, ydocentes, zdocentes, ...). The goal is to create three experimentation environments:

X – no indexes and no integrity constraints.

Y – with the standard integrity constraints (primary keys and foreign keys).

Z – with the standard integrity constraints and the extra indexes you may find convenient, you must justify the creation of each extra index; consider testing different kinds of indexes and file organisations.

METHOD

0) Start with the justification of the constraints (Y) and the extra indexes (Z).

With respect to each question below, you need to write:

- 1) the SQL query;
- 2) the answer;
- 3) the three execution plans in the three environments, their analysis and of the corresponding estimated effort;
- 4) if the execution time is measurable, a comparison of the execution times in the three environments.

QUESTIONS

1. Selection and join.

Show the codigo, designacao, ano_letivo, inscritos, tipo, and turnos for the course ‘Bases de Dados’ of the program 275.

2. Aggregation

How many class hours of each type did the program 233 planned in 2004/2005?

3. Negation

Which courses (show the code) did have occurrences planned but did not get service assigned in the year 2003/2004?

a. Use not in.

b. Use external join, and it is null.

4. Who was the professor with the most class hours for each class type in 2003/2004? Show the number and name of the professor, the kind of class, and the total number of class hours times the factor.

5. Compare the execution plans (only the environment Z) and the index sizes for the query, giving the course code, the academic year, the period, and the number of hours of the type ‘OT’ in the academic years of 2002/2003 and 2003/2004.

a. With a B-tree index on the type and academic year columns of the ZTIPOSAULA table;

b. With a bitmap index on the type and academic year columns of the ZTIPOSAULA table.

6. Select the programs (curso) with classes with all the existing types.