FACULDADE DE ENGENHARIA DA UNIVERSIDADE DO PORTO
Mestrado Integrado em Engenharia Informática e Computação

Database Technology

Assignment nr. 3

Version 1 – Student applications

NoSQL Assignment

Goals

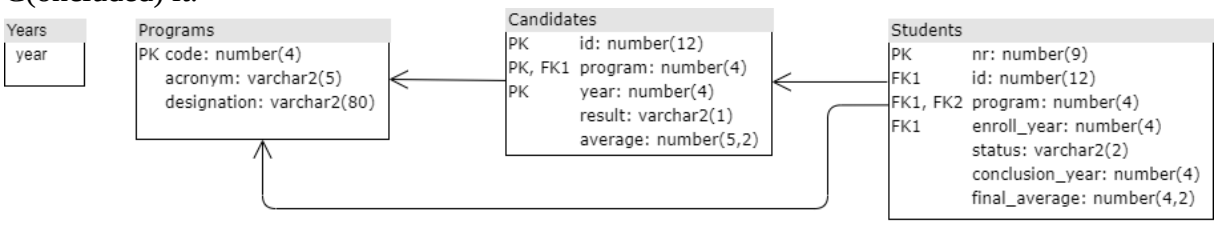
Build a small DB using two distinct NoSQL approaches, a document database (MongoDB) and a graph database (Neo4j) and compare both approaches and with the relational approach

Work groups

The assignment should be executed by a group of one or two elements.

Subject

The database to be used in this assignment is about applications, enrolments and results of students in programs, during a certain number of years. The tables are: the table programs, containing the list of programs; the table candidates with the applications; the table students with the annual enrolments of students; and the table years with the list of years with recorded enrolments. The data is about all the students that were active in 1992 or later and it was frozen in 2002 with the enrolments in 2002/2003 already done.
Before a student can be enrolled in a program in a certain academic year, his application to the program in that same year must be recorded. The student may apply for more than one program. In this phase, he is identified by the personal id (BI), as he will get a student number only in the enrolment process. With respect to each application, the intended program is known, the application outcome, and, sometimes, the application grade (calculated from the high school results). Not all the accepted candidates actually register in the program. Once enrolled, the student gets a status indicating whether he is F(ollowing) the program or he C(oncluded) it.



The tables are in Oracle user GTD12.

TASKS

The following tasks should be performed, described in a report and presented in a specific session:

1) Design a Mongo document model for the Student Applications example, explaining the decisions made. The information available under this model should be equivalent to the information in the relational database.
2) Migrate the data from the Oracle user GTD12 into the NoSQL database.
3) Design a Neo4j graph model for the Student Applications example.
4) Migrate the data from the Oracle user GTD into the Neo4j database.
5) Prepare Mongo and Neo4j queries for the following questions:
    a. Obtain the name of the program where the candidate 12147897 was enrolled.

   b. Calculate the total number of students enrolled in each program, in each year, after 1991.
   c. Obtain the BI and the student number of the students with a final grade (med_final) higher than the application grade (media).
   d. Find the average of the final grades of all the students finishing their program in a certain number of years, 5 years, 6 years, …
   e. Write a new query chosen by you.

6) Compare the Mongo, Neo4j and Oracle implementations from the viewpoints of data size, processing time, and query easiness.