University of Zurich[UZH]

Department of Informatics

University of Zürich
Department of Informatics
Binzmühlestr. 14
CH-8050 Zürich
Phone. +41 44 635 43 11
Fax +41 44 635 68 09
www.ifi.uzh.ch/dbtg

UZH, Dept. of Informatics, Binzmühlestr. 14, CH-8050 Zürich

Toni Pesut

Prof. Dr. Michael Böhlen
Professor
Phone +41 44 635 43 33
Fax +41 44 635 68 09
boehlen@ifi.uzh.ch

Zürich, June 08, 2015

BSc Thesis (17 KP)
Datenbanktechnologie

Topic: QR decomposition integration into PostrgeSQL

A DBMS offers an efficient and robust way to store and work with data. Today, an increasing number of scientists want to store and analyze scientific data inside of a DBMS, but only a few database systems, such as SciQL and SciDB, offer functionality to work with arrays and matrices and are optimized for analytic workloads. Since database systems have only limited statistical functionality, the most common way to analyze data statistically is to export the data and perform the analysis outside of the DBMS.

The goal of this project BSc project is to integrate one of the frequently used statistical operations, the QR decomposition, into PostgreSQL.

The work includes the following tasks:

1. Implement a parser extension with new following commands, where A is a table name:
   (a) "SELECT * FROM Q_QR (A);"
   (b) "SELECT * FROM R_QR (A);"
2. Implement the first part of functionality, which returns table Q, using the Gramm-Schmidt algorithm.
   (a) Suggest and implement an optimization of this part of the algorithm to reduce the number of the table scans.
3. Implement the second part of the functionality, which returns table R.
4. Run experimental analysis with different table sizes (up to 100 attributes and 1'000'000

rows)

5. Write a bachelor thesis presenting and analyzing your result (app. 50 pages).
6. Give a presentation at the DBTG group meeting (17.11.2015, 14:00-14:30).

Supervisor: Oksana Dolmatova
Start date: 12-06-2015
End date: 12-11-2015

University of Zürich
Department of Informatics

Prof. Dr. Michael Böhlen