DBT - Database Technology (4 IV) 0434 L 431

(Integrierte Veranstaltung/ Integrated Course)

Content:

The global data volume is increasing dramatically each year. Understanding how to store, process and manage these huge amounts of data efficiently is a key requirement for software engineers and dataanalysts in the modern IT world. This course will teach students both the fundamentals of data processing in traditional single-node database systems and how to scale out these techniques to huge amounts of data in large-scale, distributed environments.
The database technology course is split into two parts, each covering roughly one half of the semester. During the first part, the students learn the fundamentals of database technology for relational database systems. This includes the general architecture of a DBMS, file- and buffer management, query processing, indexing, metadata management, query optimization, locking, recovery and transaction management. In the second part, students learn the basics of parallel data processing, with a focus on largescale, distributed systems and "cloud computing". Topics include parallel relational databases, parallel no-SQL processing platforms like MapReduce, distributed data storage and retrieval – e.g., via DHTs –, techniques for distributed locking and transaction handling, multi-tenancy and software as a service, as well as modern hardware, benchmarking, and data stream processing. The course consists of a lecture and theoretical, written exercises.

Target group:

This course is the base course for **master students** with focus on database systems and information management and should be attended in the first semester of the master program.

Prerequisite:

In contrast to the introduction of database systems (MPGI5/DBS/Informationssystemse&Datenanalyse), which looks at database systems from an application programmers point of view, this class focuses on the internals of database systems. To participate, students are required to have successfully completed a Bachelor in computer science with a focus on database systems (participation in the Datenbankpraktikum, Datenbankprojekt). Knowledge of data modeling, relational algebra, and SQL as well as a very good command of Java, or possibly C/C++/C#, programming is required to participate in the course. Due to capacity reasons, the class is limited to at most 60 participants.

Registration:

Students are required to register via the DIMA course registration tool before the start of the first lecture (http://www.dima.tu-berlin.de). Within the first six weeks (**30.11.2018**) after commencement of the lecture, students will have to register for the course at **QISPOS (university examination protocol tool) and ISIS (course organization tool)** in addition to the registration at the DIMA course registration tool.

Contributions:

Prüfungsform: Portfolioprüfung
The final grade according to § 47 (2) AllgStuPO will be calculated with the faculty grading table 2.
(Die Gesamtnote gemäß § 47 (2) AllgStuPO wird nach dem Notenschlüssel 2 der Fakultät IV ermittelt.)

Studienleistung	Punkte
Assessment of 4 homework exercises with 5 pts.	20
End term test	40
Mid term test	40

Short Comment:

The lab capacity limits this course to at most 60 participants.
The module can be completed in one term.
Lectures are accompanied by individual exercises to practically rehearse the theory taught in the lectures.
The course will be given in English.

Contact persons:

Prof. Dr. Tilmann Rabl, Viktor Rosenfeld

Mon 14 – 16 in TC 006 (starts Oct 15th 2018)
Thu 10 – 12 in MA 041 (starts Oct 18th 2018) 4 SWS/6 ECTS