Database systems architecture

Lecturer
Mahmoud SAKR (Coordinator)

Course mnemonic
INFO-H417

ECTS credits
5 credits

Language(s) of instruction
English

Course period
First term

Campus
Solbosch

Course content

This course transforms you from a database user into a database programmer. No, not a programmer who connects from java or python to a database and uses it, rather a programmer who builds a database management system DBMS. From a user point of view, a DBMS is a software that stores data, so that it can later be queried in SQL. It is also expected among others that the DBMS will maintain the data security, manage the access rights, respond quickly, and handle concurrent reads and writes - all being done automatically. Database systems reached unprecedented success in abstracting away the complexities of data management and hiding them from users. Users only describe their data needs in SQL without any knowledge of how the data is physically organized.

This course explores the components of a DBMS that makes this data isolation possible. The course will cover the following topics:

> Query Execution

> Refreshing SQL and Relational Algebra

> Query Optimization

> Indexes

> Cost-based optimization

> Crash recovery

> Concurrency control

> Distributed databases

Objectives (and/or specific learning outcomes)

In contrast to a typical introductory course in database systems where one learns to design and query relational databases, the goal of this course is to get a fundamental insight into the implementation aspects of systems designed to manage and process large amounts of data. Our objective in this respect is two-fold. (1) To gain the background required to design and implement future data management and processing systems and (2) to gain an understanding of how performance of practical data management systems can be tweaked.

In particular, we take a look under the hood of relational database management systems, with a focus on query and transaction processing.

> Upon successful completion of this course, the student is expected to master the following competences :

> Translating a given SQL expression into the Relational Algebra

> Improving a relational algebra expression by, where possible, removing redundant joins in select-project-join subexpressions

> Improving a relational algebra expression by, where possible, (a) replacing cartesian products by joins ; and (b) pushing selections and projections

> Describing and being able to implement traditional secondary-memory index structures (BTrees, Hashing)

> Being able to describe and demonstrate the shortcomings of traditional index structures with respect to multi-dimensional search keys. In addition, being able to describe and implement the studied multidimensional indexes

> Given a logical query plan and given base statistics about the size and distributions of the database relations, constructing a heuristically optimal physical query plan, by estimating the sizes of the intermediate results and correspondingly comparing the possible implementations. When joins can be reordered, choosing the order with the least cost.

> Describing and being able to implement Abstract Data Types in extensible database systems

> Describe data and query distribution mechanisms, and being able to configure and run a distributed database system

Pre-requisits and co-requisits

Courses having this one as co-requisit

INFO-F439 | Advanced Methods in Bioinformatics |5 crédits, INFO-H419 | Data warehouses |5 crédits, INFO-Y099 | Multicore programming |6 crédits and MEMO-H504 | Mémoire de fin d'études en Informatique |20 crédits

Required knowledge and skills

This course assumes that participants have background on the following topics. There will be no formal check of this background at the inscription time. It is up to the participants to see and decide themselves.

> Relational databases

> SQL

> Relational Algebra

> General programming skills

Teaching method and learning activities

Cocktail of lectures, exercise sessions, and project work.

Contribution to the teaching profile

> Be capable of formulating and solving complex or open-ended technical and scientific problems by using abstraction, modeling, simulation, and multi-disciplinary analysis while satisfying the requirements of university-level research and responding to requirements, constraints, the set context and the technical, socio- economical ethical and environmental stakes—all with the purpose of obtaining concrete solutions.

> Have in-depth knowledge and understanding of a structured body of knowledge, both transversal and specialised. Be capable of autonomously and critically following current trends and advances in this body of knowledge.

> Define, plan, manage, and execute projects taking into account their objectives, the available resources and constraints; assuring the coherence and quality of the work and deliverables.

> Work efficiently with other professionals (in group, in partnership, or in competition), make decisions and develop leadership, in a variety of professional contexts, disciplines, and cultures.

> Communicate and share information in a structured manner: orally, graphically and written, in French and in one or more other languages. Communicate on scientific, technical and cultural aspects, adapting him/herself to the desired goal as well as the target audience.

References, bibliography and recommended reading

> Course notes distributed during lecture.

> Selected book chapters and papers will be recommended during lectures for further reading.

Course notes

Université virtuelle

Other information

Place(s) of teaching

Solbosch

Contact(s)

Lecturer: Mahmoud SAKR. Email: mahmoud.sakr@ulb.be

Assitant: Mariana Duarte. Email: mariana.machado.garcez.duarte@ulb.be

Evaluation method(s)

Other

Evaluation method(s) (additional information)

Project work and written exam

Determination of the mark (including the weighting of partial marks)

> Projects: 8/20

> Written exam: 12/20

Main language(s) of evaluation

English

Programmes

Programmes proposing this course at the Brussels School of Engineering

MA-IRIF | Master of science in Computer Science and Engineering | finalité Professional/unit 1 and finalité Big Data Management and Analytics (Erasmus Mundus)/unit 1

Programmes proposing this course at the faculty of Sciences

MA-INFO | Master in Computer science | finalité Professional/unit 1 and finalité Professional/unit 2