

EECS3421 Wrap Up

EECS3421 - Introduction to Database Management Systems

What is a database system?

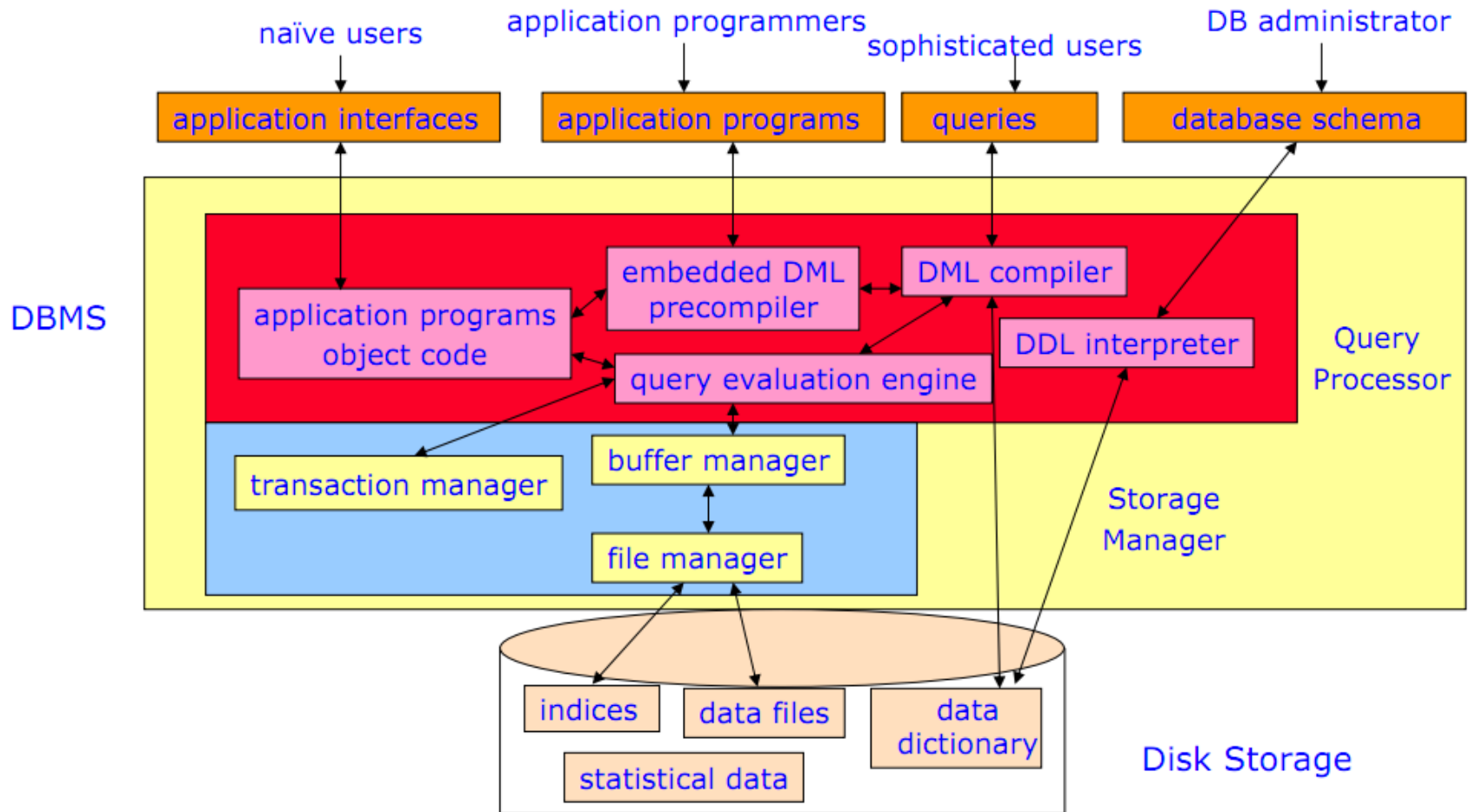
- Database: a large, integrated **collection of data**
- Models relevant aspects of reality
 - **Entities** (teams, players)
 - **Relationships** (Lionel Messi *plays for* Barcelona)
 - **Constraints** (*at least* one goalkeeper per team)
 - More recently, active components (“business logic”)
- Database Management System (DBMS): a software system designed to **store, manage, and facilitate access to** databases

Why study databases?

- Shift from computation to information
 - always true for corporate computing
 - Web made this point for personal computing
 - more and more true for scientific computing
- Need for **data science** has exploded
 - **Corporate**: retail swipe/clickstreams, “customer relationship mgmt”, “supply chain mgmt”, “data warehouses”, etc.
 - **Scientific**: digital libraries, Human Genome project, Sloan Digital Sky Survey, physical sensors, grid physics network
- A practical discipline spanning much of
 - OS, languages, theory, AI, multimedia, logic
 - Yet with a focus on real-world apps



DBMS High-level Architecture



Advantages of a DBMS

- Data independence
- Efficient data access
- Data integrity & security
- Data administration
- Concurrent access, crash recovery
- Reduced application development time
- **So why not use them always?**
 - Expensive/complicated to set up & maintain
 - Cost & complexity must be offset by need
 - General-purpose, not suited for special-purpose tasks (e.g. text search!)

EECS3421 Overview

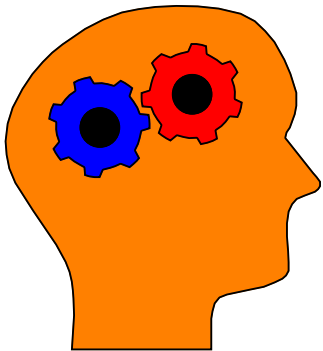
- Relational Database Modeling
 - The Relational Model & Algebra
 - Design Theory for Rel. Databases (FDs, Normal Forms)
 - High-level Database Models (E/R model, Relational Schema)
- Relational Database Querying & Programming
 - Interactive SQL (DDL, DML, Constraints, Views, Indexes)
 - Embedded SQL (Stored Procedures, Prepared Statements)
 - Database Security (SQL Injection)
- Modeling and Querying Non-relational Databases
 - Introduction to NOSQL
 - NOSQL Taxonomy by Data Model (Column, Key-Value, Document, Graph stores)

What comes next?

- Further studies at EECS@Lassonde:
 - **eeecs4411**: Database Management Systems
 - **eeecs4414**: Information Networks
 - **eeecs4415**: Big Data Systems
- If you are heading for industry:
 - Database & **data science** professionals are **in high demand**
- If you want to do research:
 - Many interesting problems ahead [The Beckman Report on Database Research, Oct 2013]
 - Scalable Big/Fast Data Infrastructures
 - Diversity in the Data Management Landscape
 - End-to-End Processing and Understanding of Data
 - Cloud Services
 - Roles of Humans in the Data Life Cycle

Summary

DB professionals are the bedrock of the information economy



Data management R&D spans a broad, fundamental branch of the science of computation

Semester goal: become an effective DBMS user (Did you?)

Thanks!

Contact:

Manos Papangelis, LAS 3050

papaggel@eecs.yorku.ca

www.eecs.yorku.ca/~papaggel/