Databases

0 Introduction

Data bases 2

Instructor: Sara Comai

sara.comai@polimi.it

"Data bases 2"
Academic Year 2019/2020

- Teacher: Prof. **Sara Comai**
  - E-mail: sara.comai@polimi.it
  - Tel. 02-2399-3611
  @Dipartimento di Elettronica Informazione e Bioingegneria, 1st floor

  - Exercise sessions:
    Prof. Daniele Braga, Pietro Pinoli, Arif Çanakoğlu, Ava Vali

Time:	
Monday	8:30-10:00
Tuesday	8:30-10:00

Textbooks

IN ITALIAN:

P. Atzeni, S. Ceri, P. Fraternali, S. Paraboschi, R. Torlone
"Basi di dati: Architetture e linee di evoluzione" (2003)

P. Atzeni, S. Ceri, P. Fraternali, S. Paraboschi, R. Torlone
"Basi di dati" (2018)

IN ENGLISH:

P. Atzeni, S. Ceri, S. Paraboschi, R. Torlone
"Database systems" - McGraw-Hill (1999)

NOW DOWNLOADABLE (http://dbbook.dia.uniroma3.it/)
Compared to the last Italian version some chapters are missing

Teaching material

- Material is available on the Beep portal

http://beep.metid.polimi.it

under "**Data bases 2**"

- Materials include slides of the lectures, exercise sessions, forum, etc.

- If your study plan has not been approved yet, you can subscribe to the course. Please specify the motivation!

App for iOS

- App to solve/check the classification of the schedules (VSR, CSR, 2PL etc.).

- Available on the App Store

"DBSA" (Data Base Schedule Analyzer)

https://itunes.apple.com/us/app/data-base-schedule-analyzer/id619821068?l=it&ls=1&mt=8

- Of course you cannot use it during the **exams: mobile devices** will be **forbidden** also to check the time.

Prerequisites

- Basics of Database systems
  - Relational model
  - SQL (SQL-92)
  - Relational algebra

  - **BOOK: http://dbbook.dia.uniroma3.it/**

Program

Two main streams:

- (Relational) Database architectures

- Advanced database systems

Study of "inside" DB technology: why?

- DBMSs provide "transparent" services:

  - So transparent that it is perfectly normal to use them ignoring many implementation details

  - So far, we have seen DBMSs as a "black box"

- So... why should we open the box?

  - Knowing how it works may help to use it better

  - Some services are provided separately

DataBase Management System — DBMS

A system (software product) capable of managing data collections that are:

- **large** ((much) larger than the central memory available on the computers that run the software)

- **persistent** (with a lifetime which is independent of single executions of the programs that access them)

- **shared** (in use by several applications at a time)

guaranteeing **reliability** (i.e. tolerance to hardware and software failures) and **privacy** (by disciplining and controlling all accesses).

Technology of DBMSs - topics

- Concurrency control

- Buffer and secondary memory management

- Reliability control

- Distributed architectures

- Physical data structures and access structures

- Query management ("optimization")

DB Evolution

Since the 70's: relational databases + SQL

Some revolutions in the 90's:
- SQL'92
- SQL'99 (triggers, object-oriented features)

And more recently:
- SQL:2003 (XML-related features)
- SQL:2006 (XQuery)
- SQL:2011 (Temporal DB)
- SQL:2016 (row pattern matching, JSON)

- Since 2005: NoSQL DBMS (no standard!)
- A single application may involve different kinds of data → diverse data models and query languages for a single application
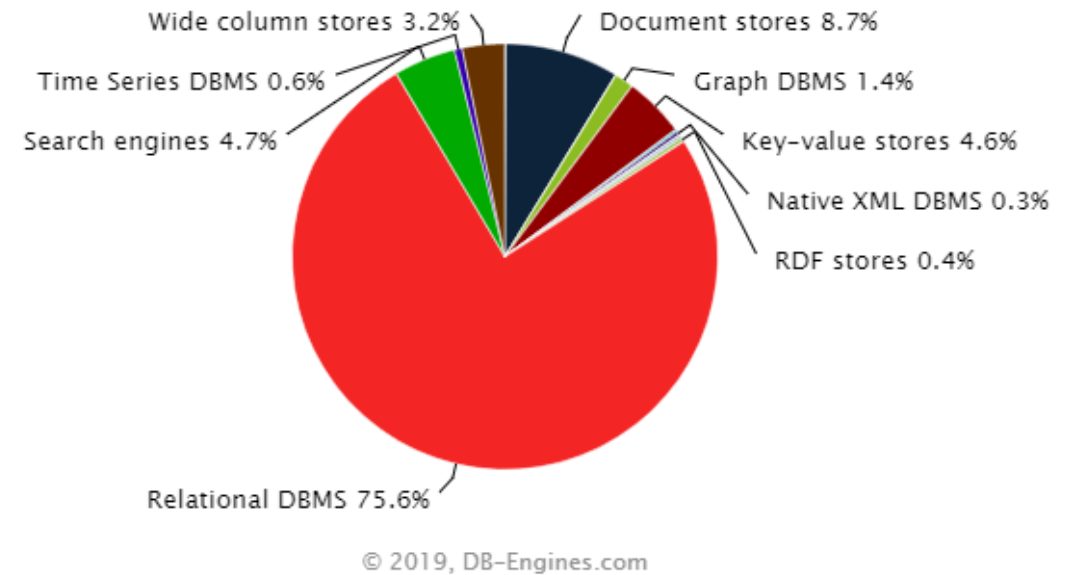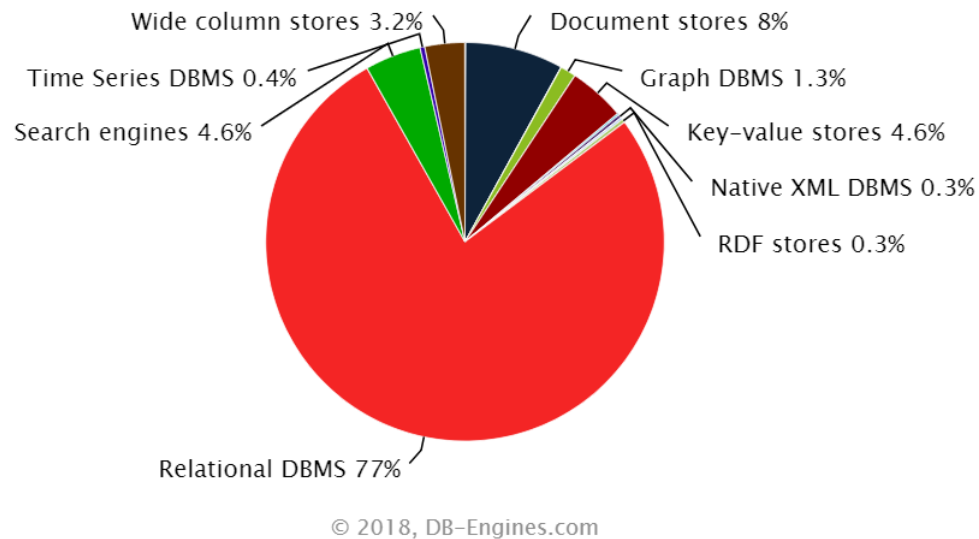
Popularity of the models

352 systems in ranking, September 2019

Rank			DBMS	Database Model	Score		
Sep 2019	Aug 2019	Sep 2018			Sep 2019	Aug 2019	Sep 2018
1.	1.	1.	Oracle ➕	Relational, Multi-model 🛈	1346.66	+7.18	+37.54
2.	2.	2.	MySQL ➕	Relational, Multi-model 🛈	1279.07	+25.39	+98.60
3.	3.	3.	Microsoft SQL Server ➕	Relational, Multi-model 🛈	1085.06	-8.12	+33.78
4.	4.	4.	PostgreSQL ➕	Relational, Multi-model 🛈	482.25	+0.91	+75.82
5.	5.	5.	MongoDB ➕	Document	410.06	+5.50	+51.27
6.	6.	6.	IBM Db2 ➕	Relational, Multi-model 🛈	171.56	-1.39	-9.50
7.	7.	7.	Elasticsearch ➕	Search engine, Multi-model 🛈	149.27	+0.19	+6.67
8.	8.	8.	Redis ➕	Key-value, Multi-model 🛈	141.90	-2.18	+0.96
9.	9.	9.	Microsoft Access	Relational	132.71	-2.63	-0.69
10.	10.	10.	Cassandra ➕	Wide column	123.40	-1.81	+3.85
11.	11.	11.	SQLite ➕	Relational	123.36	+0.65	+7.91
12.	12.	⬆13.	Splunk	Search engine	87.01	+1.12	+12.98
13.	13.	⬆14.	MariaDB ➕	Relational, Multi-model 🛈	86.07	+1.11	+15.43
14.	14.	⬆16.	Hive ➕	Relational	83.10	+1.30	+23.46
15.	15.	⬇12.	Teradata ➕	Relational, Multi-model 🛈	76.97	+0.32	-0.42

Ranking from http://db-engines.com/en/ranking

Popularity of the models

Ranking scores per category in percent, August 2018

Ranking scores per category in percent, September 2019

http://db-engines.com/en/ranking

Exam

- The exam consists in a written verification covering all the topics of the course

  - **Exercises on the whole program, possibly with related theoretical questions.**

  4 main topics:
  1. DB architecture/technologies (concurrency control etc.)
  2. Trigger
  3. Xquery
  4. Physical DB/Query optimization

Exam rules

- During the exam:

  - <u>No books, notes, electronic devices </u>are allowed

  - Cheating policies: it is forbidden to communicate with other students. Who is surprised to talk, is asked to leave the classroom. This applies to both "extremes" of communication: both speaker and listener.

- After the exam

  - <u>No oral exams </u>will be done

Exam rules

- Positive marks can be **rejected**

  - This can be done through the online system; usually there are 5 days after the insertion of the mark into the system to reject it

  - REMARK: when the exam is repeated (it is sufficient to sit down and see the text of the exam), the **previous mark is lost!**

Academic calendar and DB2 exams

- 1st semester: ends before Christmas

  1st examination session: 2 DB2 exams
  - Between January 9, 2020 and February 22, 2020

  2nd examination session: 2 DB2 exams
  - Between June 11, 2017 and July 31, 2017

  3rd examination session: 1 DB2 exam
  - Last week of August – first two weeks of September (calendar not available yet)

- **Do not leave during the examination sessions!!**
  - Requests for extra-exams will be rejected

> **17 January 2020**
> **7 February 2020**