

DB 2

Summer 2022

Torsten Grust
Universität Tübingen, Germany

1 | Welcome all...

... to this course which is all about digging in the mud of database management system (DBMS) internals.

We will try to understand how DBMSs

1. **organize primary and secondary memory** (RAM and SSD/HDD) to store and query tables of data,
2. use “**data maps**” (**indexes**) to navigate huge tables,
3. internally **represent, optimize, and evaluate complex queries** over tables, and
4. **manage concurrent access to data** while avoiding inconsistencies and confusion.

But Why Would You Do That?

Insights into the DBMS kernel and its inner workings can explain, for example, ...

- why **query evaluation** takes (much) longer than expected (we are talking hours 🐢 vs. msec 😊 here),
- whether **database growth** (by factor $n > 1$) will slow down updates and/or queries by factor n (or n^2 or ...),
- whether the host's **resources** (CPU, cache, RAM, storage) are used effectively or if bottlenecks exist,
- how (known) **algorithms and data structures** hold up in the presence of huge inputs.

SQL is the Secret Key

I assume that you have working knowledge of **SQL**,¹ the “*Intergalactic Dataspeak*” used to manipulate and query tabular (or: relational) DBMSs.

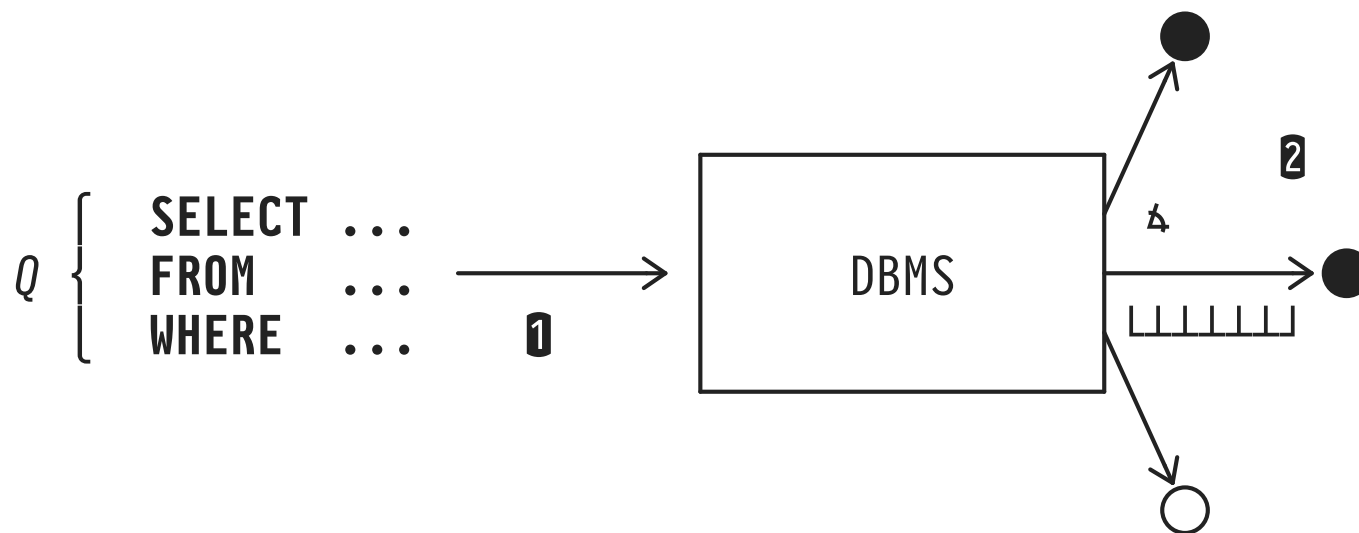
We will **submit a variety of SQL queries** Q to DBMSs and observe

- which components of the DBMS engage in processing Q ,
- how DBMS-internal data structures and algorithms contribute to execute Q efficiently.

¹ As taught in our courses “DB 1” or “Advanced SQL”, for example.

Sending “SQL Probes”

- 1 Submit SQL probe query Q to DBMS, then...
- 2 ... observe/interpret DBMS response.



Thankfully, DBMSs provide several hooks and facilities that support the observation of their operation.

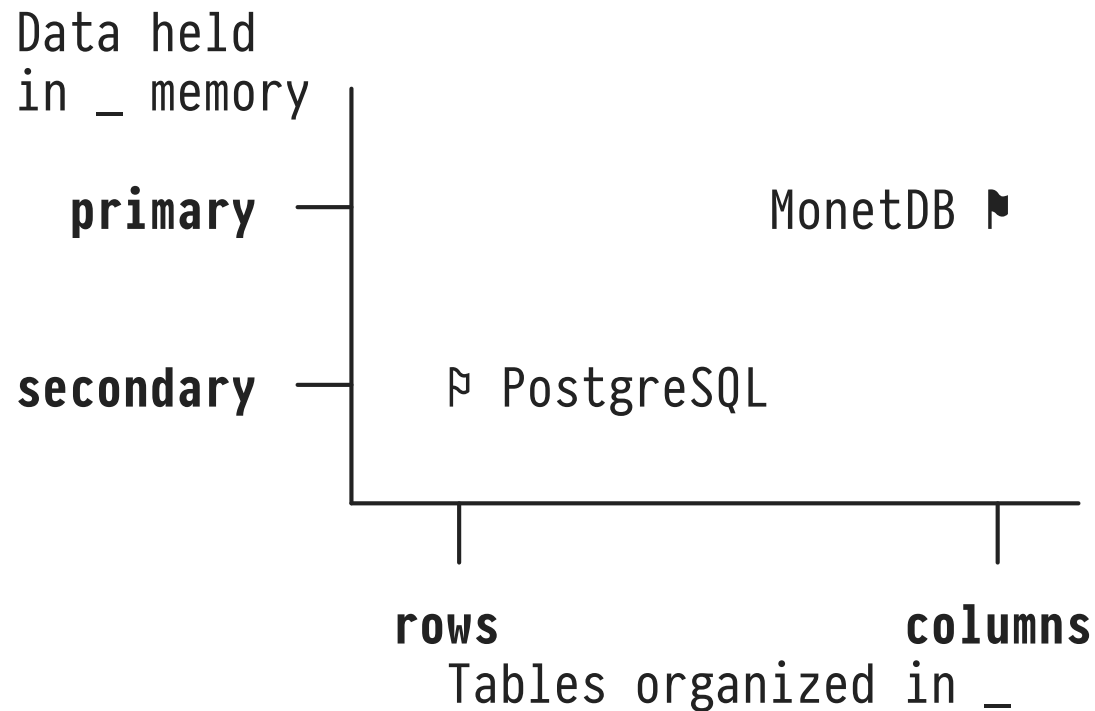
EXPLAIN Yourself, DBMS!

We will make extensive use of **EXPLAIN** facilities that act like a “DBMS X-ray.” Instead of Q , submit **EXPLAIN Q** to reveal

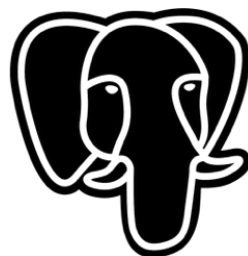
- the **evaluation order** chosen to process the expressions/statements in Q ,
- the **indexes** used to access tables,
- the **cost** and **resources** (# of I/O operations, memory buffers, CPU time) expended to process Q ,
- the **algorithms** selected to access tables, manipulate rows, and compute Q 's results.

2 : Two Ends of a Spectrum: PostgreSQL and MonetDB

There is **one relational data model** but a **variety of DBMSs** that implement it. We will focus on two DBMSs, **PostgreSQL** and **MonetDB**, whose internals deviate significantly:



PostgreSQL (postgresql.org)



- Row-wise table organization (☰)
- Data held in blocks on secondary storage (e.g., SSD/HDD), database size virtually unlimited (32 TB in single table)
- Extensible database kernel (types, operations), support for a rich SQL:2011/SQL:2016 dialect
- Actively developed since 1986 (derived from Ingres)
- In this course: **PostgreSQL 12.10**
- *Open source*, available on macOS 🍏, Linux 🐧, Windows 🪟

MonetDB (monetdb.org, developed at CWI Amsterdam)



- Column-wise table organization (▣)
- Data completely mapped into primary memory for processing
—once hot database size » RAM size, MonetDB struggles
- Optimized for CPU performance
- Actively used in database research since 1993 until today
(*"the column store pioneers"*)
- In this course: **MonetDB 5** (v11.35.19 or newer should do)
- *Open source*, available on macOS 🍏, Linux 🐧, Windows 🪟

On the Bare Metal

Occasionally we will discuss/develop short **program fragments in C** to understand how DBMSs—MonetDB, in particular—process data internally.

There is no need to be a C whizz, but we will encounter:

- dynamic allocation of memory (`malloc(3)`)
- array types (`τ[]`), array access, array processing
- pointer types (`τ*`), pointer operations
- control flow (conditionals, loops)
- UNIX system calls (`open(2)`, `mmap(2)`, ...).



Torsten Grust?

Time Frame	Affiliation/Position
1989–1994	Diploma in Computer Science, TU Clausthal
1994–1999	Promotion (PhD), U Konstanz
2000	<i>Visiting Researcher</i> , IBM (USA)
2000–2004	Habilitation, U Konstanz
2004–2005	Professor Database Systems, TU Clausthal
2005–2008	Professor Database Systems, TU München
since 2008	Professor Database Systems, U Tübingen

- E-Mail: Torsten.Grust@uni-tuebingen.de
- Twitter: [@Teggy](https://twitter.com/Teggy) (*Professor, likes database systems, programming languages, and SC Freiburg ツ*)
- WSI, Sand 13, Room B318 (home office!)

Administrivia

Weekday/Time	Slot	Room
Monday, 10:15–11:45	Lecture	Sand 6/7, F119
Tuesday, 10:15–11:45	Lecture	Sand 6/7, F119
Thursday, 14:15–15:45	Tutorial	Sand 6/7, F119



-  **No** lectures/tutorials on
 - Thu, April 21 (tutorials start on April 28)
 - Thu, May 26 (Ascension Day)
 - Mon/Tue/Thu, June 6/7/9 (Whitsun break)
 - Mon/Tue/Thu, June 13/14/16 (SIGMOD 2022)
 -  lecture videos, assignment will be handed out

Administrivia

End-Term Exam

- 90-min **written exam** on Monday, July 25, 8:30–10:00 in N10+N11 (Morgenstelle).
- You may bring a DIN A4 double-sided *cheat sheet*.
- Passing earns you 9 ECTS.
- Score $\geq \frac{2}{3}$ of the overall assignment points to be admitted to the exam and earn bonus points in the end-term exam.

Weekly Assignments & Tutorial Sessions

- We will distribute, collect, and grade **weekly assignments** (Friday→Friday) via Github .
- You work on these in **teams of two**. Hand-in again via .

Organized and run by **Denis Hirn**:



- E-Mail: Denis.Hirn@uni-tuebingen.de
- WSI, Sand 13, Room B312

Assignments start once we have collected the first batch of interesting material, probably on April 28.

Lecture Videos, Slides, and Pieces of Code

- These **slides** (PDF) and **code fragments** (SQL, C, ...) will be uploaded to a Github  repository:

github.com/DBatUTuebingen/db2-ss22 

- For the 2020 edition of the course, I have produced **lecture videos**:
 - 83 videos, \approx 30-min fragments.
 - Playlist on YouTube : tinyurl.com/DB2-2020
 -  Since 2020, the course has moved on—material may be added, superseded, or shuffled.
 - We do aim to make your/our time in F119 worthwhile.

Forum

During this summer semester, the **DB2 forum** is *the* course hub:

forum-db.informatik.uni-tuebingen.de/c/ss22-db2 

- ⚠ **Registration** (mandatory) and announcements
- ❓ Questions and answers (do *not* post complete solutions)
- ☁️ Download additional code examples (SQL, MAL, and C)
- 💬 General discussion
- 🕒 Quick turnaround (responses often within minutes)

Course Homepage

db.inf.uni-tuebingen.de/teaching/DB2SS2022.html 

- **Organizational matters**

Curriculum. General announcements regarding the lecture, exams, or dates. Please surf by regularly. Thank you!

- **Contact information**

Turn to the forum first. But feel free to send e-mail if you seek specific help/need to discuss personal issues with us.

Material

This course is *not* based on a single textbook but instead draws from

- a variety of scientific papers,
- textbook excerpts,
- blog and mailing list postings, [Stack Exchange](#) Q&As,²
- SQL references/standards,
- DBMS docs for [PostgreSQL](#) (★★★★★) and [MonetDB](#) (★★★★☆),
- experience, and best practices.

² All of dba.stackexchange.com/questions/tagged/{postgresql,monetdb} are worth a look.

Material (on Index Design and Usage)



- To the point, clear, and actionable advice on how to design “data maps”—or: **indexes**—and how to recognize whether a query can (not) benefit from an index.
- See use-the-index-luke.com [free HTML version] and sql-performance-explained.com [PDF 9.95€, Paperback 29.95€].