Kick-Off

Course Goals

Methods:

- Practice to read scientific papers
- Learn how to write a scientific paper
- Learn how to give a scientific talk
- Learn how to discuss scientific literature and give feedback

Content:

1. Fundamental approaches to query execution
2. Efficient execution of queries and operators
3. Selected modern query engines

Prerequisites

Hard prerequisite:

- Introduction to Fundamentals of Databases (IN0008) or equivalent

Beneficial previous knowledge:

- Database Systems on Modern CPU Architectures (IN2118)
- Most courses from Prof. Neumann, Prof. Giceva, Prof. Kemper, and Prof. Leis

Expected Work

During the course each of you will create:

- A research report **draft** (5-8 pages + references):
  - summarizing the assigned topic and discussing it
  - also in context of other papers discussed in the seminar
- A presentation (13-15 minutes)
- 2 peer reviews for your fellow students ($<$1 page each)
- A **final** research report (5-8 pages + references):
- Meaningful contribution to the panel discussion in your presentations session

During the course each of you will read 8-10 papers:

- your assigned paper
- the paper also presented in your slot (to spark the discussion)
- 2-3 fundamental papers for each of the three topic blocks

Grading

Rough estimate of grading contributions:

- $\approx 45\%$ Report
- $\approx 25\%$ Presentation
- $\approx 10\%$ Peer reviews
- $\approx 20\%$ Active discussion

This is subject to change!

Preliminary Topics (1)

Fundamental approaches to query execution

must MonetDB/X100: Hyper-Pipelining Query Execution.

must Efficiently compiling efficient query plans for modern hardware

must Everything you always wanted to know about compiled and vectorized queries but were afraid to ask

topic Relaxed operator fusion for in-memory databases: Making compilation, vectorization, and prefetching work together at last

topic Incremental Fusion: Unifying Compiled and Vectorized Query Execution

topic Tidy Tuples and Flying Start: fast compilation and fast execution of relational queries in Umbra

topic Building advanced SQL analytics from low-level plan operators

topic Excalibur: A Virtual Machine for Adaptive Fine-grained JIT-Compiled Query Execution based on VOILA

...

Preliminary Topics (2)

Efficient execution of queries and operators

must Morsel-driven parallelism: a NUMA-aware query evaluation framework for the many-core age

must Micro adaptivity in vectorwise

topic Efficient processing of window functions in analytical SQL queries

topic These Rows Are Made for Sorting and That's Just What We'll Do

topic A practical approach to groupjoin and nested aggregates

topic The 3D hash join: Building on non-unique join attributes

topic A scalable and generic approach to range joins

topic High-Performance Query Processing with NVMe Arrays: Spilling without Killing Performance

topic Robust External Hash Aggregation in the Solid State Age

...

Preliminary Topics (3)

Selected modern query engines

must Duckdb: an embeddable analytical database

must Composable Data Management: An Execution Overview

topic Photon: A fast query engine for lakehouse systems

topic Query processing on tensor computation runtimes

topic Apache Arrow DataFusion: A Fast, Embeddable, Modular Analytic Query Engine

topic HetExchange: Encapsulating heterogeneous CPU-GPU parallelism in JIT compiled engines

topic Designing an open framework for query optimization and compilation

...

Timeline

Preliminary timeline:

- Thu 24.04.2025 introduction lecture 01 | **submit** topic preferences
- ~~Thu 01.05.2025~~ Holiday
- Thu 08.05.2025 introduction lecture 02
- Thu 15.05.2025 introduction lecture 03
- Thu 05.06.2025 presentation session 01
- ~~Thu 29.05.2025~~ Holiday
- Thu 05.06.2025 presentation session 02
- Thu 12.06.2025 presentation session 03 **submit** report draft
- ~~Thu 19.06.2025~~ Holiday
- Thu 26.06.2025 $\emptyset$
- Thu 03.07.2025 presentation session 04 | **submit** peer reviews
- Thu 10.07.2025 presentation session 05
- Thu 17.07.2025 presentation session 06
- Thu 24.07.2025 presentation session 07 | **submit** final report
- **Attendance to all sessions in presence is mandatory**

Organization

- Attendance to all sessions in presence is mandatory
- You have to write your paper in LaTeX using our template

Contact

- Website: http://db.in.tum.de/teaching/ss25/seminarModernDatabaseSystems
- i3mqe@db.cit.tum.de