Database Engine Development Lab
SoSe 2015

Part II

Alexander Frömmgen, Alejandro Buchmann
Project Schedule

21.04. (0) Intro, query language, performance-benchmark

28.04. (1) Store data on disc / in RAM, execute selection

12.05. (2) Execute projection, join, sort

26.05. (3) Performance benchmark

09.06. (4) Optimizer / choose most suitable execution plan

23.06. (5) Advanced stuff

14.07. (6) Advanced stuff

04.08. (7) Final demonstration
Next Steps

- Import data ([example1.zip](#))
- Store data on disc (maybe in your own format?, make assumptions, e.g. max string length)
- Read data from disc into main memory

- Implement *selection* query/operator (feel free to already implement a filter, e.g. only data from a certain experiment)
- Provide the result as datastructure or simply print it
- How long does a selection take?

Do not assume a fix data schema!
There will be more files/additional columns in the next weeks!
Next Steps

These steps imply design decisions regarding:

- Logical data structures
- Physical data/storage structures
- Buffer management
- Query processing
Blueprint of the Solution

Query language for data retrieval

Performance benchmark

Parse expression
Generate execution plans
Choose most suitable execution plan
Execute plan (projection, selection, join, aggregation…)

RAM
CPU

Store data on disk / in RAM

Inter- and intra-query parallelization

Disk

CSV

Import data

Please allow the configuration of the maximum RAM consumption!
Example

- **Input**
  > SELECT * FROM simDb

- **Output**
  > Spalte A | Spalte B | Spalte C
   wert     |  wert    |  wert
   ...     |  ...    |  ...

- **Alternativ**

```scala
// open DB
val db = new SimDB("myDB");

// read data cube
val ds = new ExperimentDBSourceDataCube(db)

// make query
// val ds2 = ds.filter().equals("Experiment", 320).avg("Node", "Time").rotate("Statistic").orderBy("Time")

// materialize
val result = ds.evaluate()
println(result.get(0))
```
What’s next?

- Meeting in **A313**,

- You will
  - present your results

- We will give you
  - feedback on your results & implementation idea

Be prepared to present your results!
Questions?