

Advanced Database Systems: Part I – Exercise

Matthias Lanzinger

Institute of Logic and Computation
Databases & Artificial Intelligence Group

March 2019

Table of Contents

- 1 Expectations
- 2 Technical Aspects
- 3 Important Dates & Questions

Table of Contents

1 Expectations

2 Technical Aspects

3 Important Dates & Questions

Submissions I

For the first exercise sheet you will submit a report in TUWEL as a *single pdf file*.

- **Absolutely no handwritten submission.**
- We recommend using \LaTeX but you are free to use whatever you want.
- Show your work and explain your answers.
Example of what not to do: 1a) The answer is 5.
- Feel free to include screenshots / graphics. However, there is a 20MB file size limit and you may have to compress your pictures.

Submissions II

Some questions are purposely open-ended. The intention is to give you room to experiment and deepen your understanding.

- **Evaluation is based on making a serious attempt and demonstrating understanding.**

You can get credits even for a solution that is not technically correct!

- Make sure to *communicate your arguments* well.
- You should be able to competently discuss your approach at the exercise interviews.
- Don't forget to actually answer the question.

Plagiarism

Plagiarism is passing off the work of others as your own.

- Plagiarism is a serious offense.
- Exercises are optional.
You don't need to submit solutions for everything.
- Your understanding will be tested at the exercise interviews.
- You will need to understand the same material for the exam.
- **You should still discuss ideas with your colleagues.**

Understand Your Submission

- Grading of exercises is based on your understanding of the material.
- Correct answers in your report do not automatically guarantee any credits!
- Every group member has to be able to demonstrate understanding of *all* of your submission.

The TUWEL Forum

The *TUWEL forum* is a place for you to ask for clarifications and share information (e.g., helpful tools) with your colleagues.

- If possible ask your questions on the TUWEL forum instead of writing us an email.
- Feel free to answer your colleagues questions.
- **Do not post (partial) solutions!**

Table of Contents

- 1 Expectations
- 2 Technical Aspects
- 3 Important Dates & Questions

Writing Your Report

- **Again: No handwritten reports.**
- Collaborative Editors:
Overleaf, Google Docs, ...
Share your recommendations with your colleagues.
- Some tips:
 - In \LaTeX , use `verbatim` blocks for queries and query plans.
 - Dia or <https://www.draw.io/> are convenient tools for drawing simple diagrams.
 - Minify your images if your report is too big (e.g., <https://compressor.io/>).

The Exercise Server

For exercise 5, use the Postgres instance on `bordo.dbai.tuwien.ac.at`. (You should have received an email with your account information).

- Supported usage is the Postgres command line client via ssh.
- On Windows you can use PuTTY:
`https://www.chiark.greenend.org.uk/~sgtatham/putty/latest.html`
- You can transfer files via scp/sftp.
Filezilla and WinSCP are good graphical interfaces for sftp.
`sftp://bordo.dbai.tuwien.ac.at`

The Exercise Server II

You only need to use the exercise server for your reported results.

- Be aware that the server might become overloaded close to the submission deadline.
- Feel free to experiment on local installations.

Local results can be very different than on the server!

SSD vs HDD, server version, memory, configuration, ...

Postgres

From the shell use `psql` to open the Postgres client.

- `\d` schema and table information.
- `\i` to read commands from a file.
- `\o` to write output to a file.
- `\a` toggles aligned output (e.g., for JSON export).
- `\h` for help.
- Tab auto-completion can be very helpful.

Query Plans

Postgres query plans play an important role in some of the exercises. Here is some further material to understand them:

- How to read query plans:
<https://www.postgresql.org/docs/9.6/using-explain.html>
- More detailed information on how plans are generated:
<https://www.postgresql.org/docs/9.6/planner-optimizer.html>

Query Plan Visualization

Complex query plans may be easier to read when visualized:

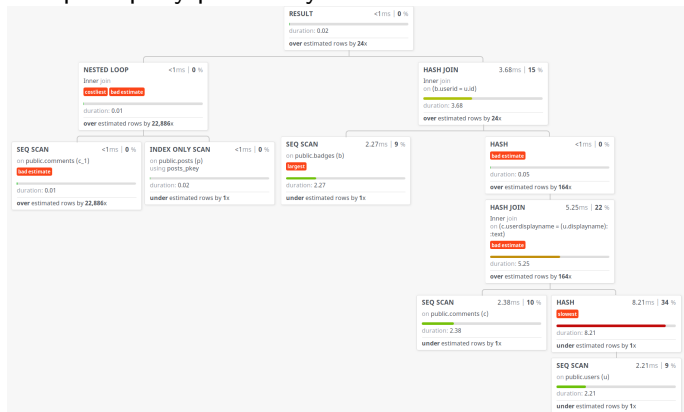
QUERY PLAN

```
-----
Result (cost=5227.89..5781.89 rows=825 width=675) (actual time=16.788..21.057 rows=34 loops=1)
  One-Time Filter: $1
  InitPlan 1 (returns $1)
    -> Nested Loop (cost=0.29..8117.06 rows=22886 width=0) (actual time=0.024..0.024 rows=1 loops=1)
      -> Seq Scan on comments c_1 (cost=0.00..948.86 rows=22886 width=4) (actual time=0.004..0.004 rows=1)
      -> Index Only Scan using posts_pkey on posts p (cost=0.29..0.31 rows=1 width=4) (actual time=0.018..0.018 rows=1)
          Index Cond: (id = c_1.postid)
          Heap Fetches: 0
    -> Hash Join (cost=5227.89..5781.89 rows=825 width=675) (actual time=16.762..21.022 rows=34 loops=1)
        Hash Cond: (b.userid = u.id)
        -> Seq Scan on badges b (cost=0.00..454.73 rows=24273 width=32) (actual time=0.003..2.090 rows=24273)
        -> Hash (cost=5221.11..5221.11 rows=491 width=643) (actual time=15.579..15.579 rows=3 loops=1)
            Buckets: 1024 Batches: 1 Memory Usage: 10kB
            -> Hash Join (cost=1573.01..5221.11 rows=491 width=643) (actual time=15.532..15.570 rows=3 loops=1)
                Hash Cond: (c.userdisplayname = (u.displayname)::text)
                -> Seq Scan on comments c (cost=0.00..948.86 rows=22886 width=228) (actual time=0.002..2.090 rows=24273)
                -> Hash (cost=616.45..616.45 rows=14445 width=415) (actual time=8.547..8.547 rows=14445)
                    Buckets: 16384 Batches: 2 Memory Usage: 1985kB
                    -> Seq Scan on users u (cost=0.00..616.45 rows=14445 width=415) (actual time=0.002..0.002 rows=1)

```

Query Plan Visualization

Complex query plans may be easier to read when visualized:



(With <https://github.com/AlexTatiyants/pev>)

Export Query Plans as JSON

To export a JSON query plan to a file you can use the following command sequence in `psql`.

You will still need to remove the first and last line to make it valid JSON.

```
\a
\o plan.json
EXPLAIN (ANALYZE, FORMAT JSON, VERBOSE, BUFFERS ON)
  SELECT ... ;
\o
```

Table of Contents

① Expectations

② Technical Aspects

③ Important Dates & Questions

Deadlines

- Upload your solutions by **08.04.2019 12:00 (lunchtime)**.
- Register for an exercise interview (as a group!) by **09.04.2019 23:55**.
- Be punctual for your exercise interview.

Question Sessions

There are two **voluntary** sessions for you to ask questions in person. These are intended for technical help, clarifications and to help you if you're stuck.

Not for checking your solutions.

- 03.04. 09:00-10:00: Informatiklabor Zelda
- 03.04. 15:00-16:00: Informatiklabor Zelda

Informatiklabor Zelda is at Favoritenstraße, staircase 4, 2nd floor.

Contacting Us

- TUWEL forum
We don't check any other forums / chats.
- Email: adbs@dbai.tuwien.ac.at
(Please stop writing to individuals)

Your Questions.