Seminar: Implementation Techniques for Main Memory Database Systems

Kickoff Meeting

Prof. Dr. Jana Giceva
Jan Böttcher, M.Sc.
Dominik Durner, M.Sc.
Philipp Fent, M.Sc.
**Michael Freitag, M.Sc.**
Maximilian Schüle, M.Sc.

July 13, 2020
Overview

Weekly Meeting
- Monday, 16:00 - 17:30, starting October 12, 2020
- Room MI 02.09.014
- 2 presentations per meeting
- There will be an attendance log

Required Work
- Seminar paper (≤ 5 pages)
- Sample implementation (C++)
- Presentation (20 minutes + 10 minutes discussion)
- Moderate one discussion (act as the "devil’s advocate", you should pair up for this)
Organization

Registration through matching system
- Register for the seminar on https://matching.in.tum.de!

After matching: Check in with the supervisor for your preferred topic
1. Check in **soon after matching** for paper recommendations (preferences considered FCFS)
2. Check in when rough structure is planned
3. Check in when first draft is ready

Due Dates
- Structure: ca. 4 weeks prior to presentation date
- Presentation slides: 1 week prior to presentation date
- Seminar paper and finished implementation: 2 weeks after presentation date
# Preliminary Topic List

<table>
<thead>
<tr>
<th>Topic</th>
<th>Supervisor</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Scalable and Robust Latches for Database Systems (1)</td>
<td>Jan Böttcher</td>
</tr>
<tr>
<td>• Scalable and Robust Latches for Database Systems (2)</td>
<td></td>
</tr>
<tr>
<td>• Towards Scalable Dataframe Systems / When sweet and cute isn’t enough anymore: Solving scalability issues in Python Pandas with Grizzly</td>
<td>Dominik Durner</td>
</tr>
<tr>
<td>• PolarFS: An Ultra-low Latency and Failure Resilient Distributed File System for Shared Storage Cloud Database</td>
<td></td>
</tr>
<tr>
<td>• BB-Tree: A practical and efficient main-memory index structure for multidimensional workloads</td>
<td>Philipp Fent</td>
</tr>
<tr>
<td>• Interpolation-friendly B-trees: Bridging the Gap Between Algorithmic and Learned Indexes</td>
<td></td>
</tr>
<tr>
<td>• External Merge Sort for Top-K Queries</td>
<td></td>
</tr>
<tr>
<td>• Updateable HyperLogLog Sketches</td>
<td>Michael Freitag</td>
</tr>
<tr>
<td>• Leapfrog Triejoin: A Simple, Worst-Case Optimal Join Algorithm</td>
<td></td>
</tr>
<tr>
<td>• White-box Compression: Learning and Exploiting Compact Table Representations</td>
<td></td>
</tr>
<tr>
<td>• HetExchange: Encapsulating heterogeneous CPU-GPU parallelism in JIT compiled engines <strong>(GPU required)</strong></td>
<td>Maximilian Schüle</td>
</tr>
<tr>
<td>• DB4ML - An In-Memory Database Machine Learning Support</td>
<td></td>
</tr>
</tbody>
</table>
https://db.in.tum.de/teaching/ws2021/seminarHauptseicherdbs/