HyPer

Technische Universität München

Data Blocks: Hybrid OLTP and OLAP on Compressed Storage using both Vectorization and Compilation

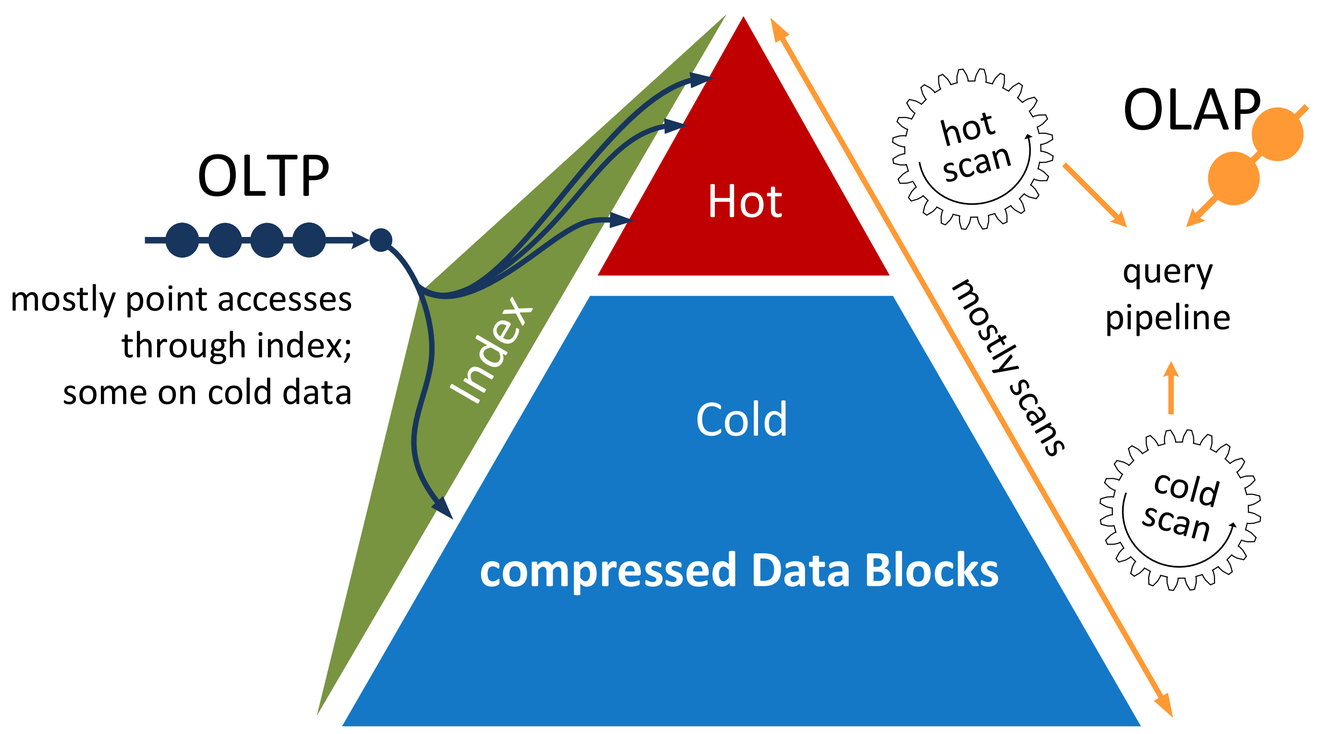
Harald Lang[1], Tobias Mühlbauer[2], Florian Funke[3],
Peter Boncz[4], Thomas Neumann[1], Alfons Kemper[1]

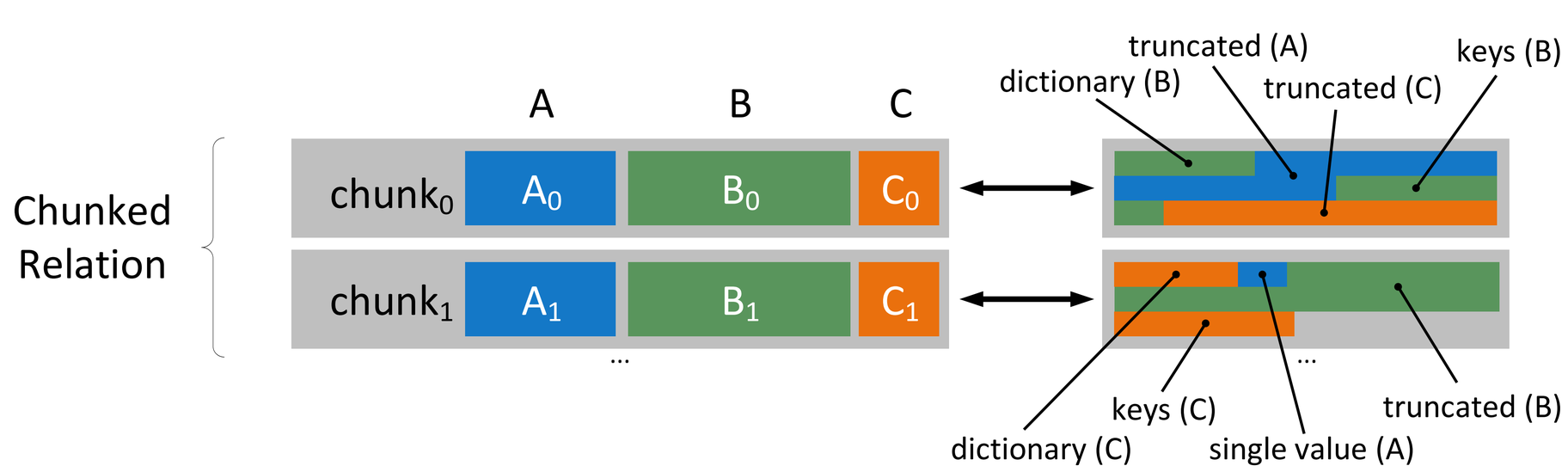
[1] Technical University of Munich, [2] Tableau Software, [3] Snowflake Computing, [4] Centrum Wiskunde & Informatica

Reducing the memory footprint in hybrid in-memory database systems

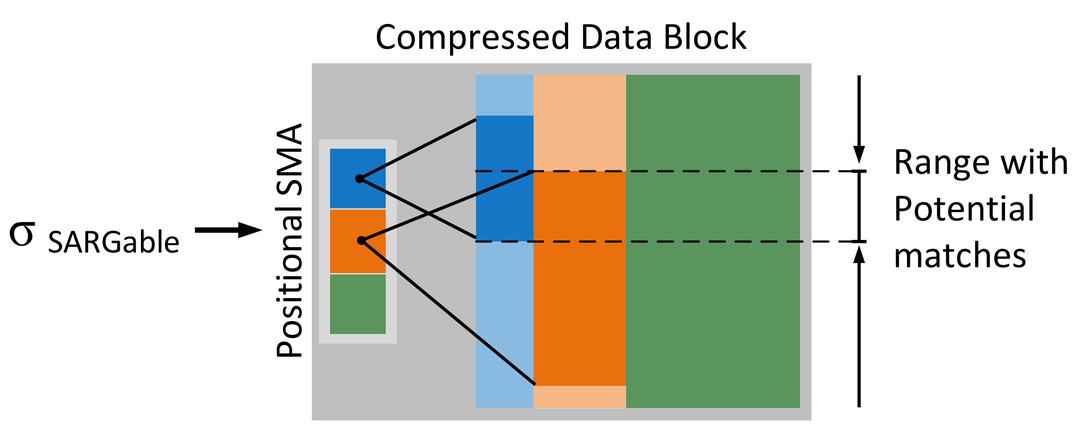
- Most data is considered as **cold data**
- **Save memory through compression**
- **Retain high transaction throughput** *and* **query performance**

- **Lightweight compression schemes**
- Attribute values remain **byte-addressable**
- **Fast scans** *and* **fast point-accesses**
- **Optimal compression scheme is chosen** for each column in each block





Positional SMA: Intra-block indexing



- **Narrow the scan range** within a block
- Lookup table where each entry contains a range with potential matches
- For *n* byte values, the table consists of *n* x 256 entries
- Only the **most significant non-zero byte is considered**

Integration of vectorization and JIT-compilation

- Huge variety of physical Data Block memory layouts
- Necessitates **interpretation**
- **Vectorized scans**
- Find and extract matches