Two Birds With One Stone: Designing a Hybrid Cloud Storage Engine for HTAP

Tobias Schmidt, Dominik Durner, Viktor Leis, Thomas Neumann
Technische Universität München
tobias.schmidt@in.tum.de

Hybrid Transactional and Analytical Processing (HTAP)

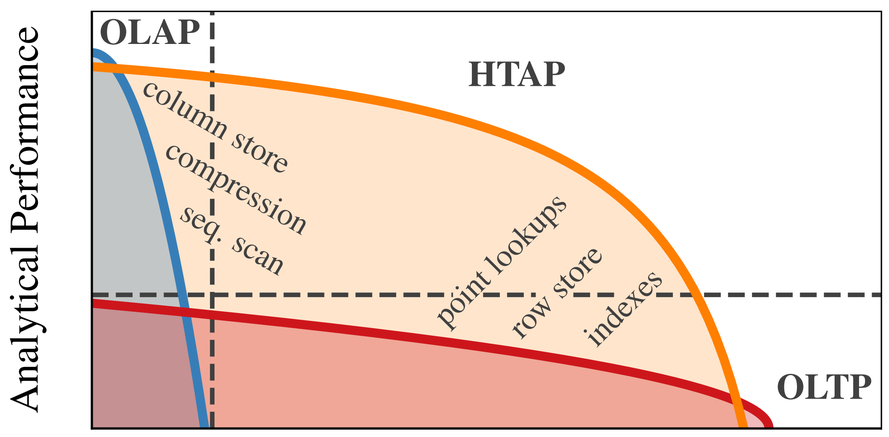
Problem: Systems either optimize for OLTP or OLAP workloads.

OLTP: row-oriented storage, index structures, point lookups, small updates → *fail to process large data sets efficiently.*

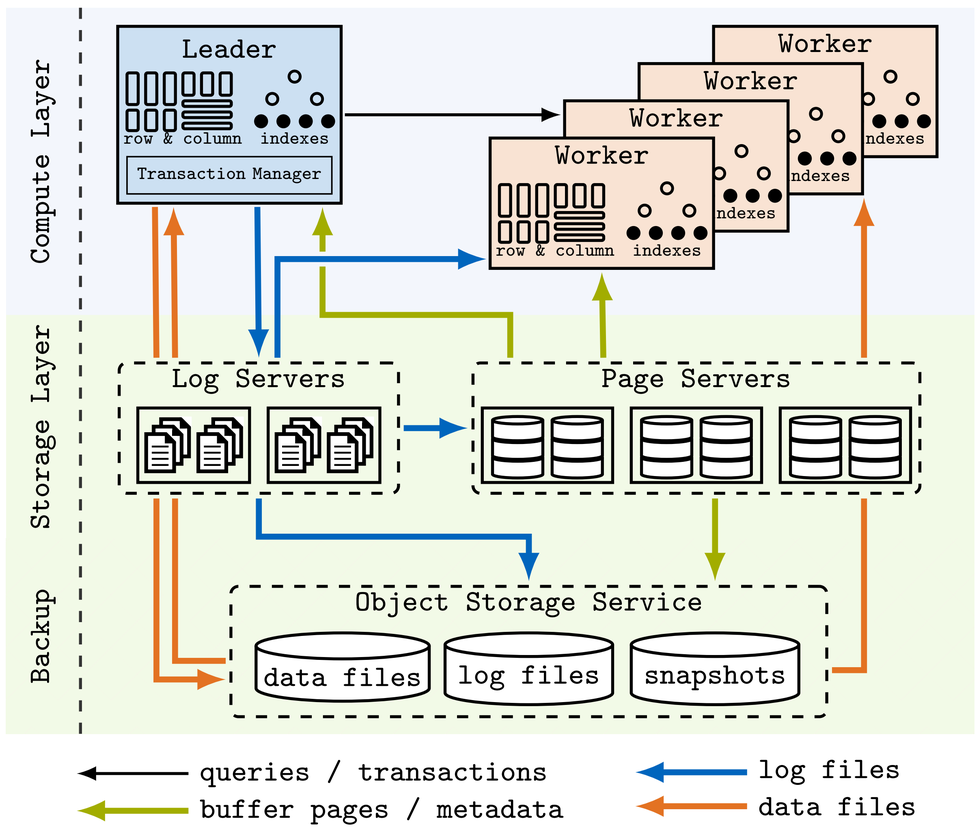
OLAP: column-oriented storage, vectorized processing, compression, sequential scans → *fail to provide high transaction rates.*

Colibri: Storage engine design for processing hybrid workloads efficiently on solid-state drives and cloud object stores.

Performance Characteristics



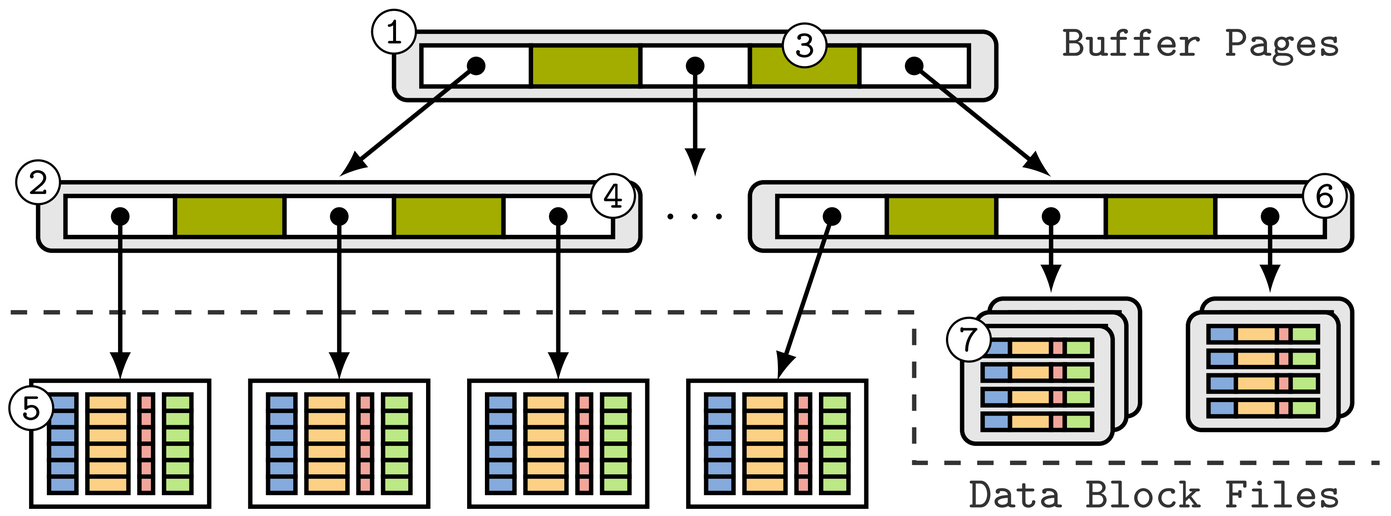
Colibri: a Cloud-Native HTAP Database Architecture



Design Principles:

(1) Separation of hot and cold data
(2) Exploiting high-bandwidth storage devices
(3) Minimize the number of log entries
(4) Bypassing page and log servers

Hybrid Row-Column Store:

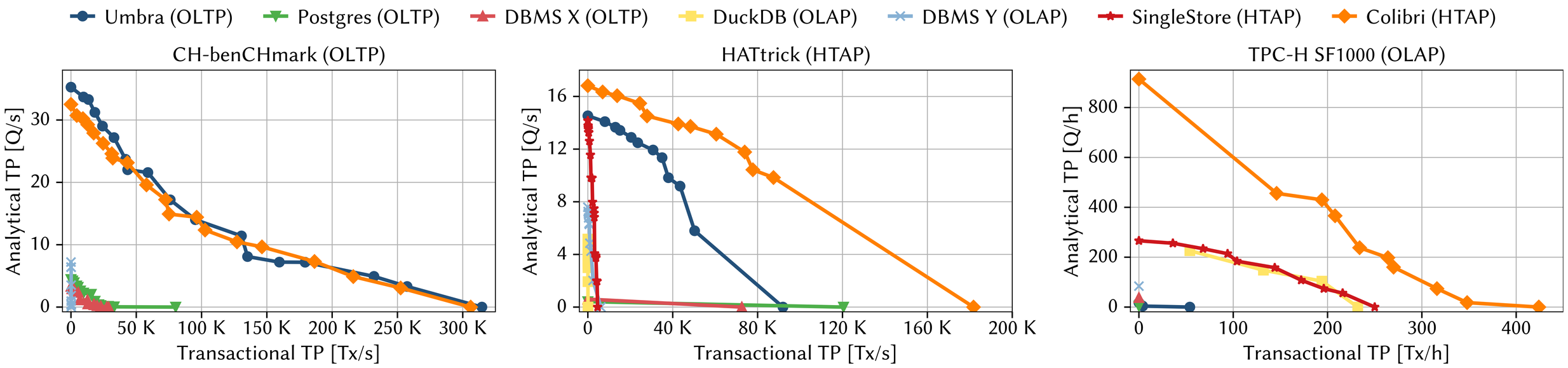


(1) B$^+$-tree indexes (5) compressed data blocks (columnar format) and (7) uncompressed data pages (PAX format) based on (3) row ids.

Core Features:

(a) Bandwidth-optimized scan architecture
(b) Lightweight multi-version concurrency control protocol
(c) Bypassing the write-ahead log
(d) Custom compressed data format
(e) Swip-based buffer manager
(f) Optimistic lock coupling

Evaluation: Analytical Performance vs. Transactional Throughput



Colibri's hybrid column-row store bridges the gap between analytical and transactional systems.

VLDB2024, 50th International Conference on Very Large Databases, Guangzhou, China