

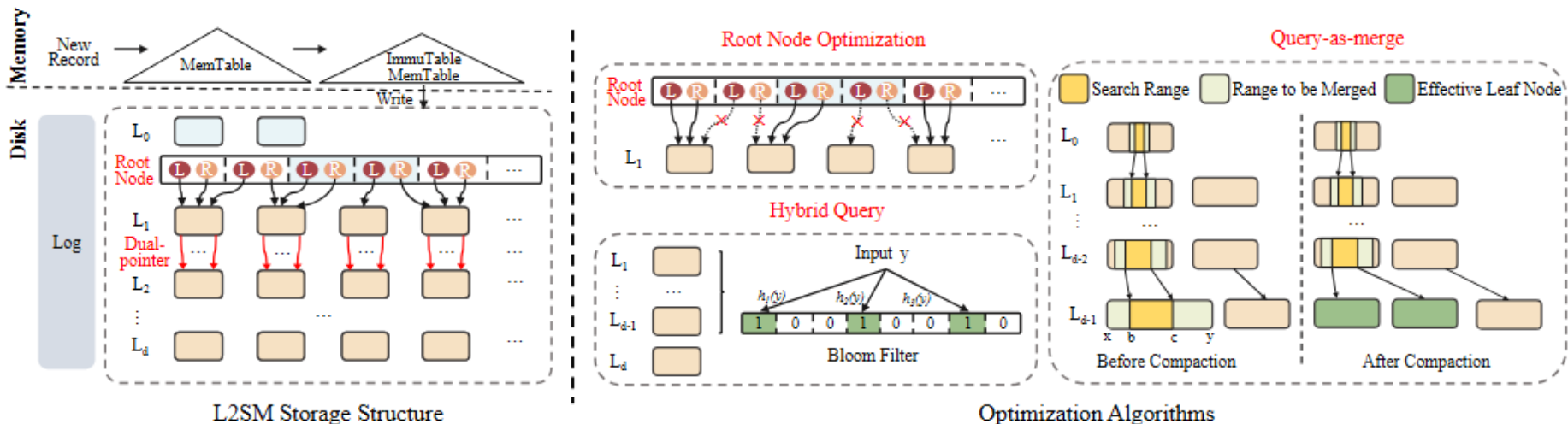
# L2SM: A Query-optimized Linked LSM-tree for HTAP Workloads

**Xiaoyue FENG, Dashan WEI, Chaopeng GUO, Jie SONG**

Frontiers of Computer Science, DOI: [10.1007/s11704-024-40553-0](https://doi.org/10.1007/s11704-024-40553-0)

# Problems & Ideas

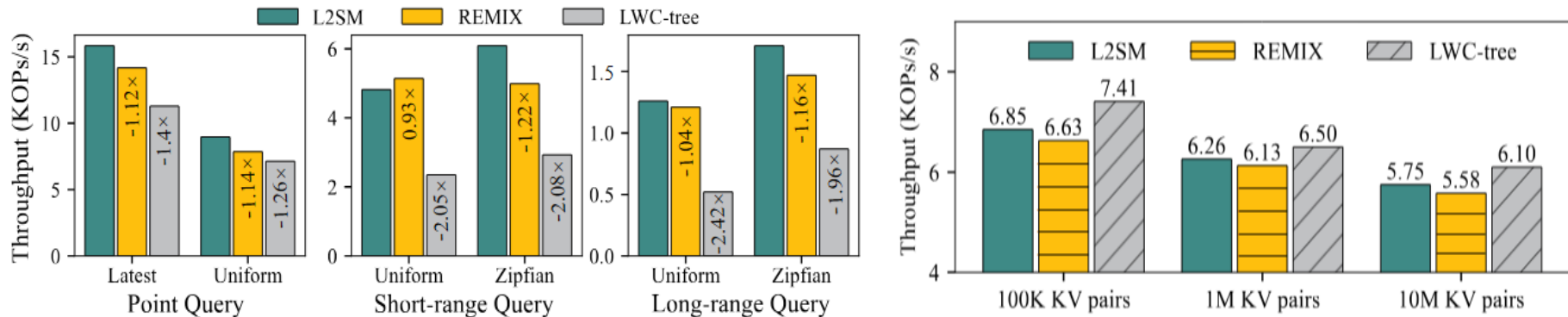
- Problems of LSM-tree optimization approaches:
  - Existing optimization approaches fail to provide high-performance range queries while guaranteeing efficient write performance, which makes the database less capable of handling AP tasks when processing HTAP workloads.
- Our idea:
  - a root-node-linked, cascading storage structure called L2SM (Linked-LSM tree) to improve the query efficiency and reduce the memory overhead.



The root node can help efficiently determine the required data in  $L_0$  and the cascading storage structure with dual-pointer allows to directly find all data within the defined range in each level. They improve the performance of processing range queries, which is the main query type of AP tasks.

# Main Contributions

- Contributions:
  - We propose a cascading dual-pointer storage structure with a root node, named L2SM, which establishes index pointers for data pairs to enhance query performance;
  - We propose three optimization algorithms for the L2SM structure, including root node global optimization, hybrid query, and query-as-merge;
  - Our experiments demonstrate that L2SM shows stable write performance and excellent query ability compared to LSM-tree, REMIX tree, and LWC-tree.



Query and write performance comparison results for three structures. Left: Throughput of three structures for processing three types of queries under different distributions; Right: Throughput of three structures for processing random write.