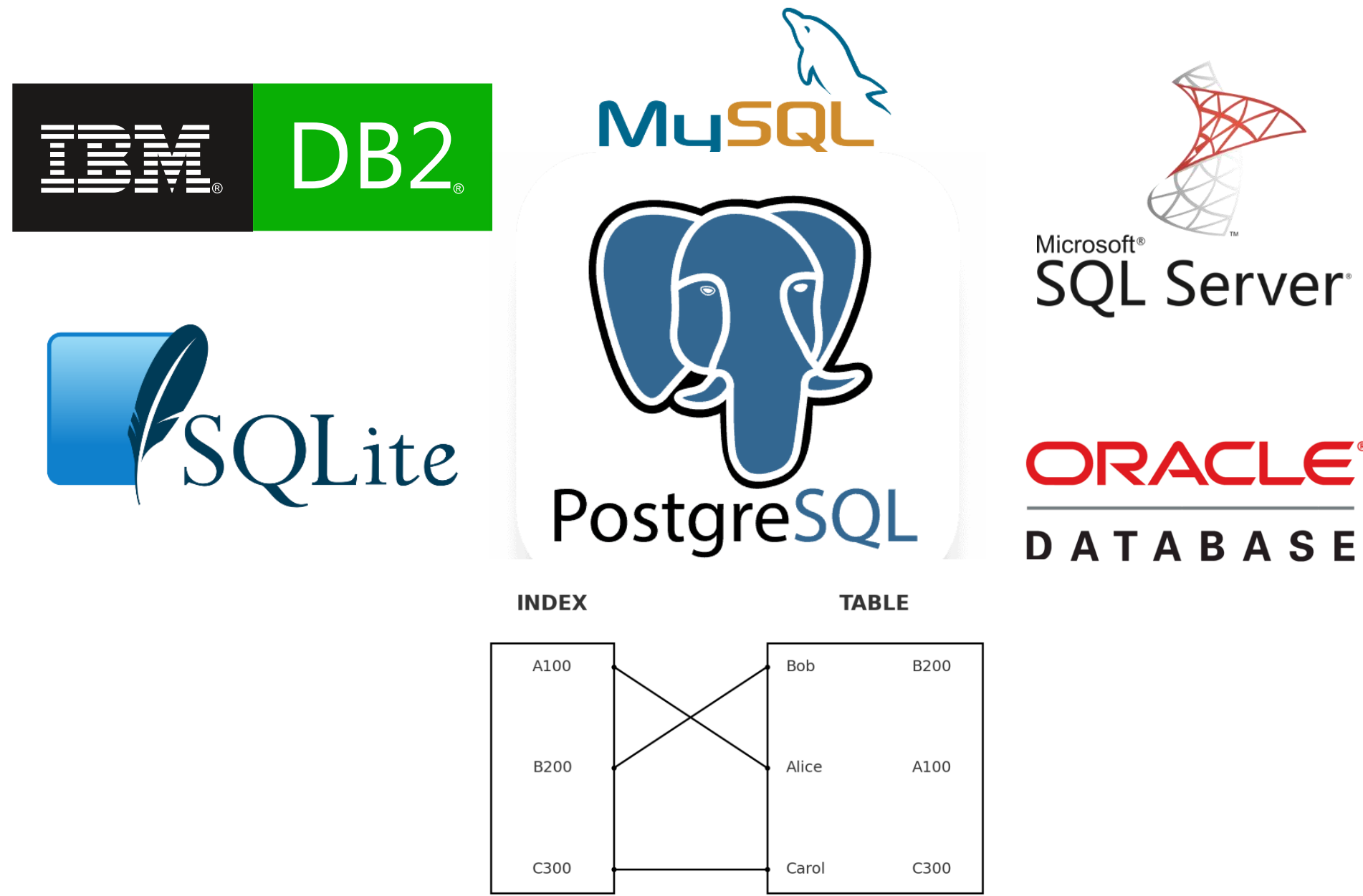
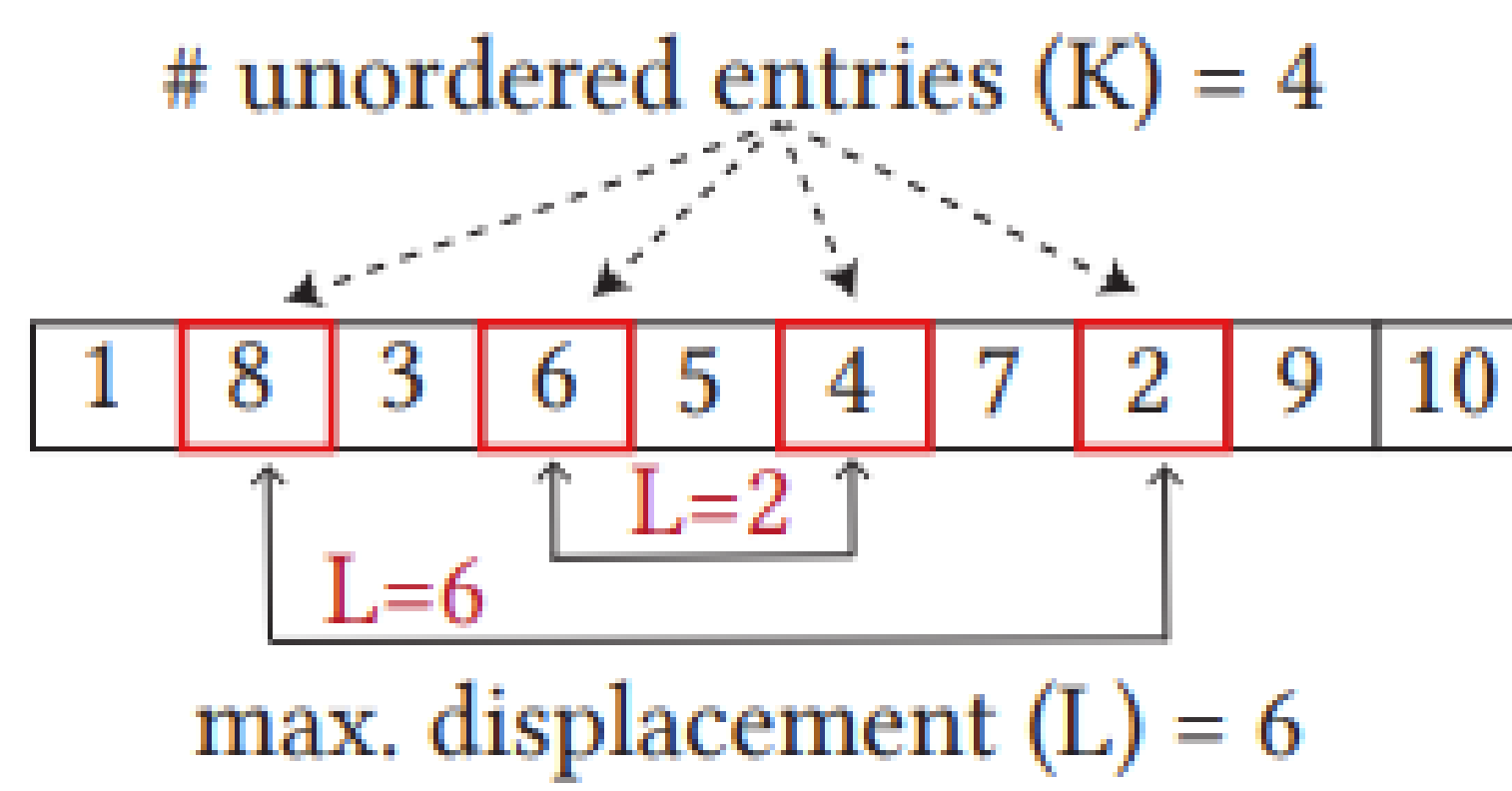


Indexes in Data Systems



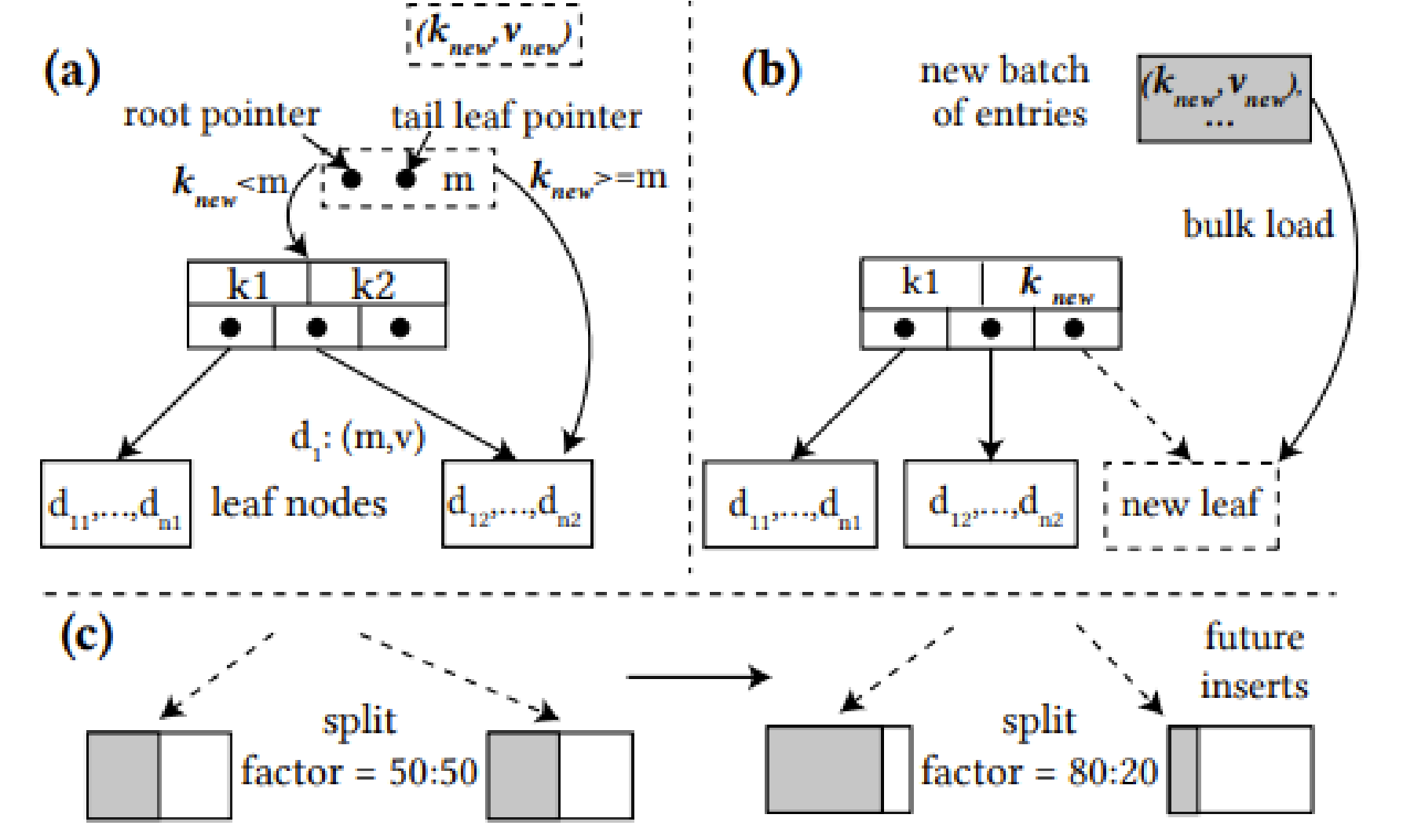
Indexes in databases help facilitate efficient queries

K-L Sortedness Metric



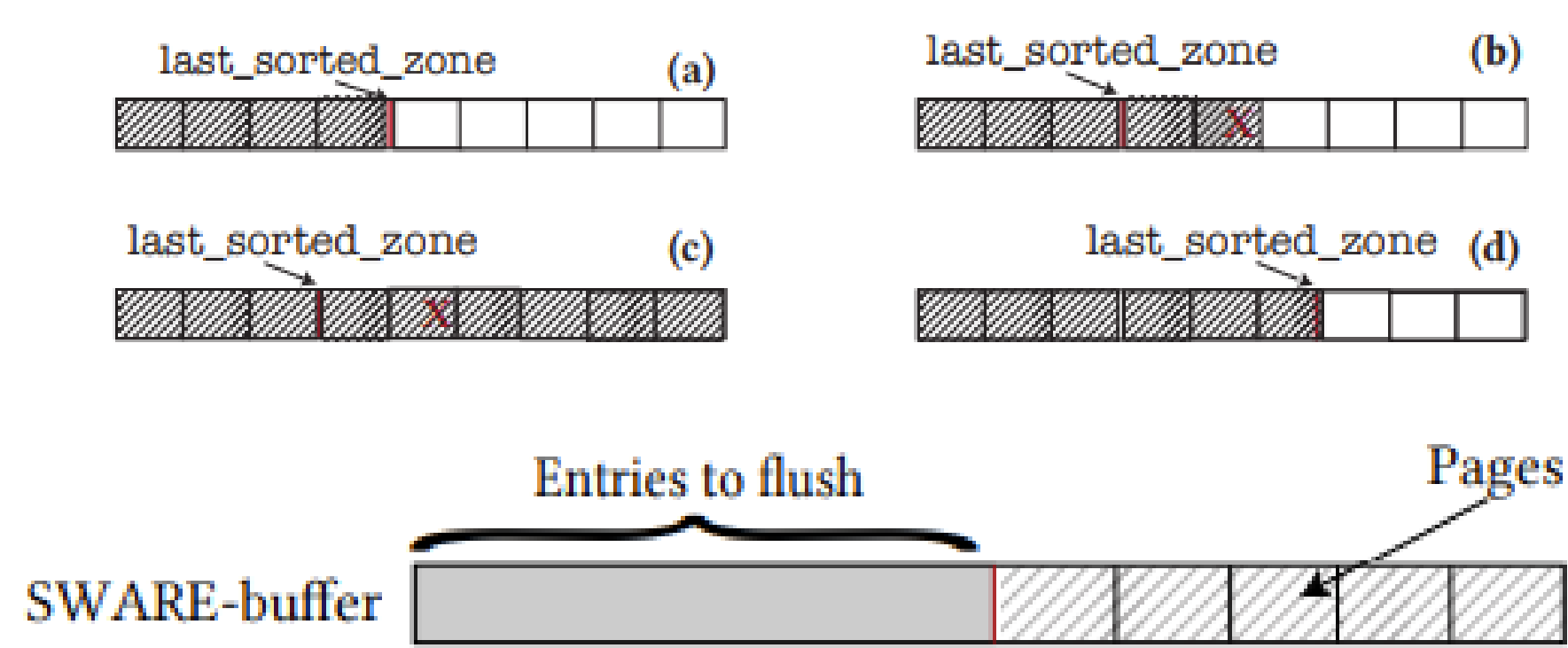
The K-L metric is used to measure the level of data sortedness in a data set
 $K = \#$ out of order entries
 $L = \max$ displacement of an out of order entry

SWARE



An SA-B⁺-tree uses a state-of-the-art buffer to opportunistically bulk load data, decreasing insertion and construction costs

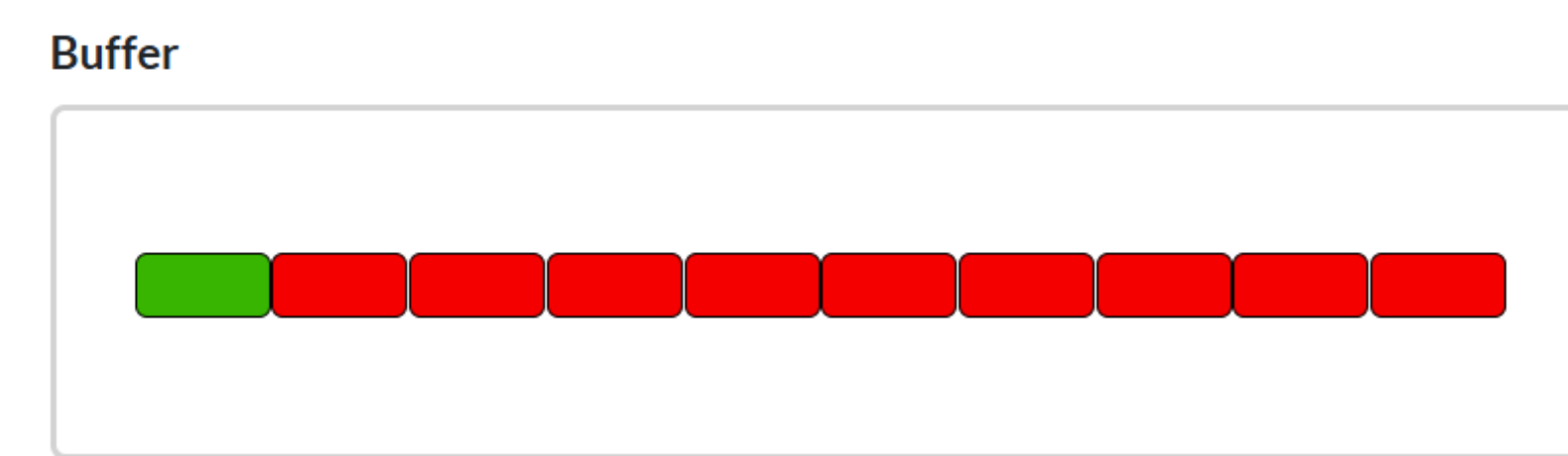
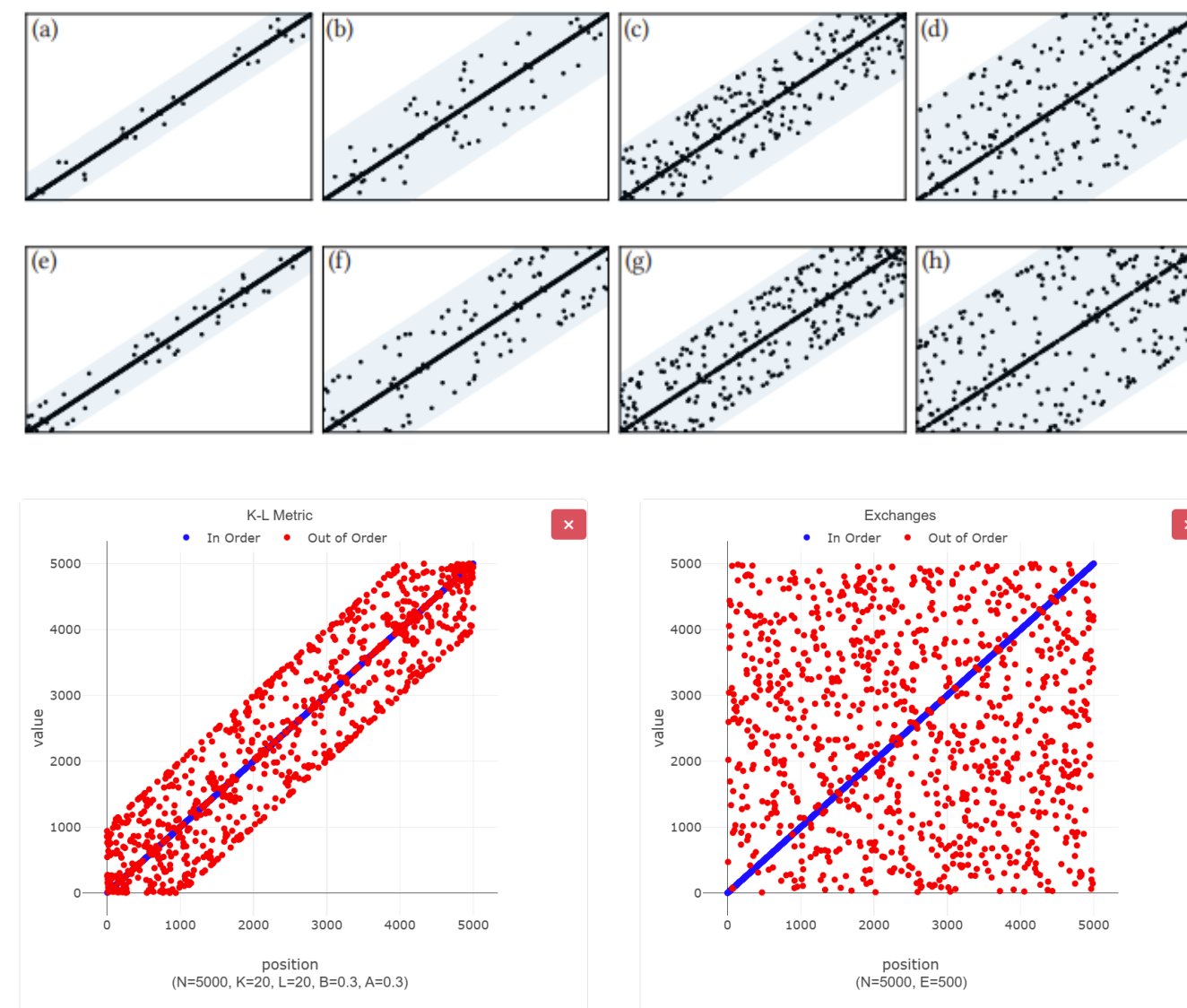
SWARE Buffer



- SWARE buffer batches entries
- Tracks sorted section and opportunistically flushes data to maximize bulk loading
- Buffer induces overhead during lookups

Methods

Adapted BoDS data generator in Javascript that uses the K-L-sortedness metric to visualize the data



Implemented SWARE buffer and B⁺-tree in JavaScript along with a visual of data ingestion



Demonstration

Input Panel

Workload

N (# OF ELEMENTS) 5000 K (% OUT-OF-ORDER) 10 L (% MAX OUT-OF-ORDERNESS) 10

Sortedness Distribution

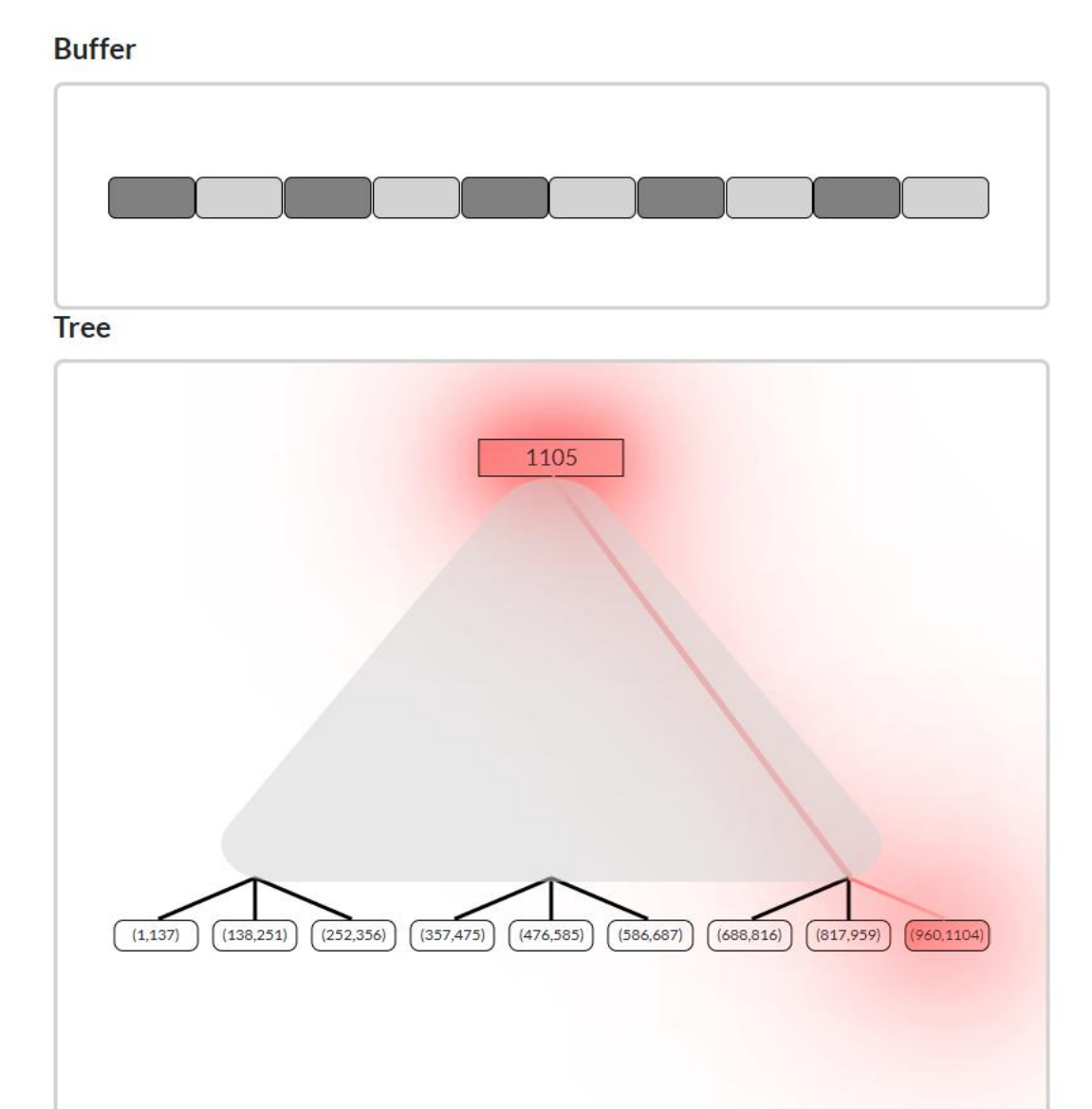
α (ALPHA VALUE) 0.3 β (BETA VALUE) 0.3

Run

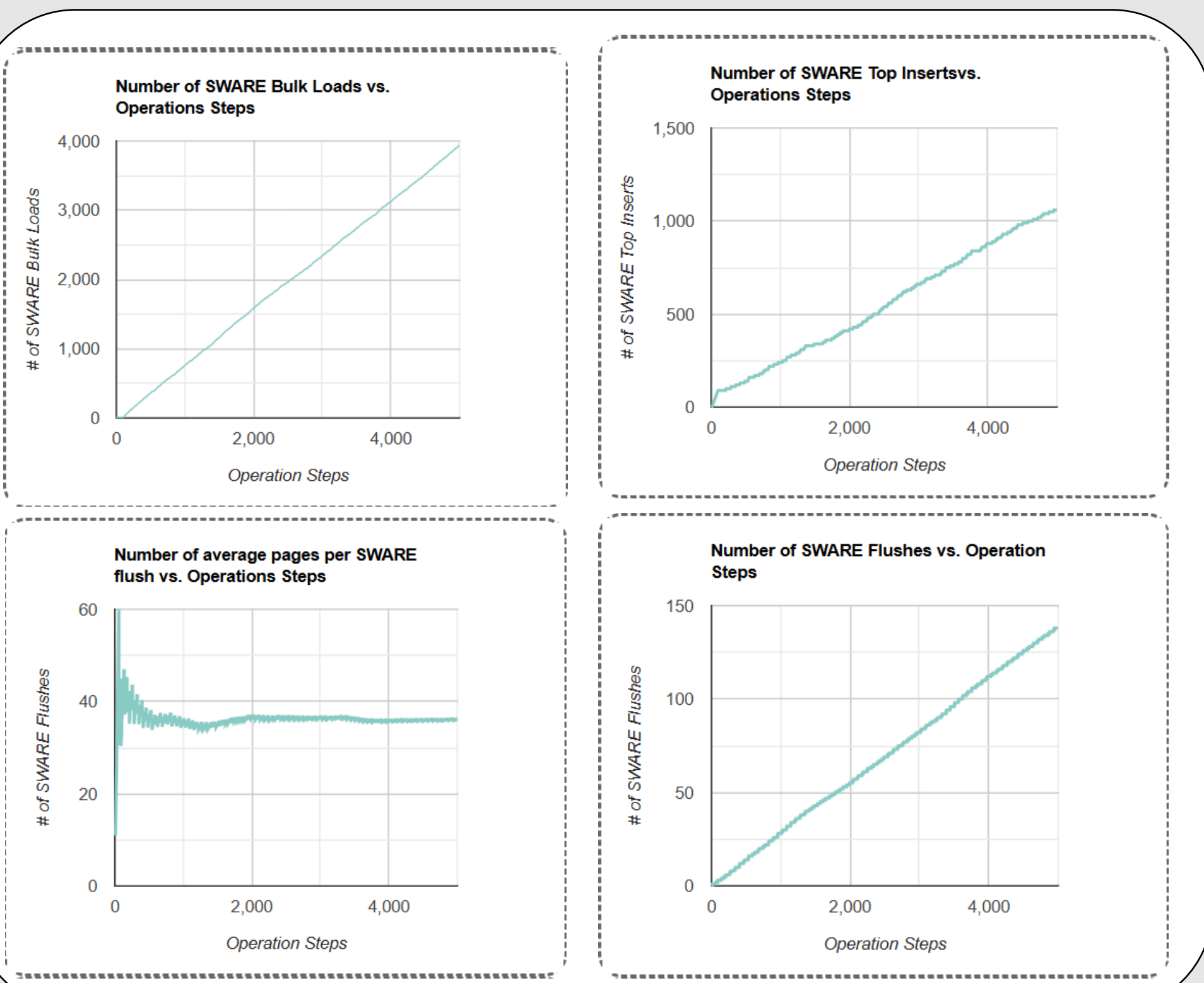
Data Visualization



Emulation Panel

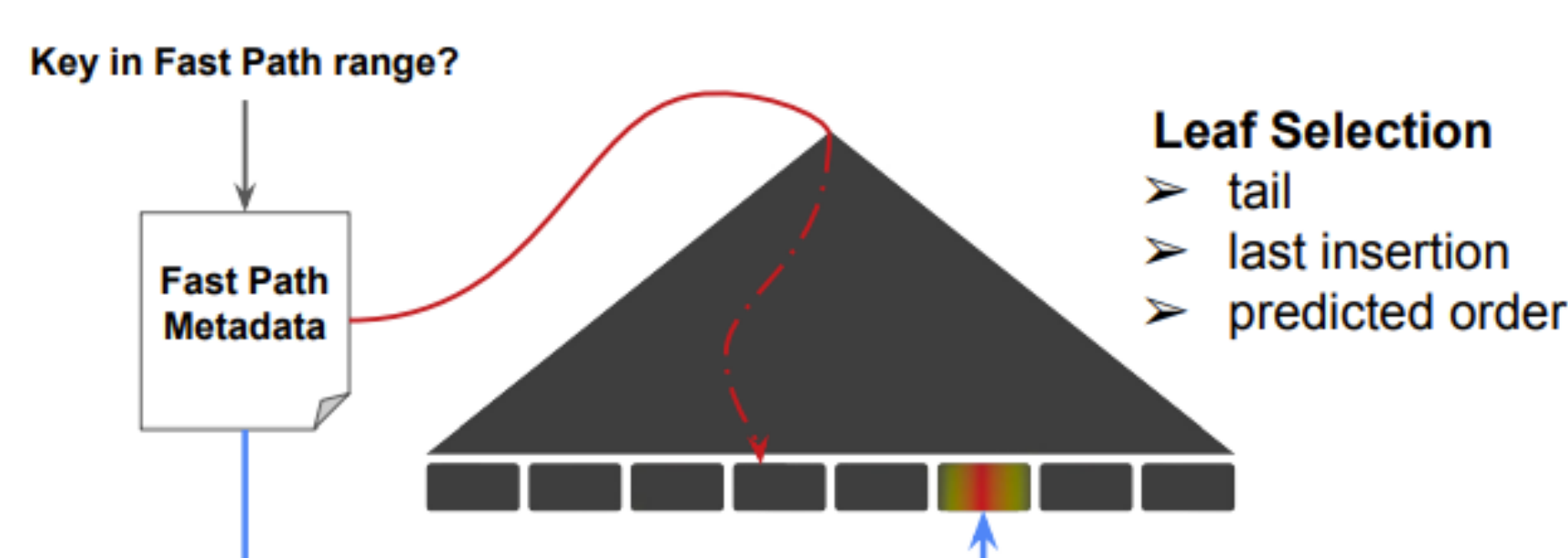


Observations



Future Work

Expand the demonstration to other sortedness-adaptive index designs like tail-B⁺-tree, lil-B⁺-tree and QuIT



References

Raman, A., Karatsenidis, K., Sarkar, S., Olma, M., & Athanassoulis, M. (2022, September). BoDS: A Benchmark on Data Sortedness. In *Technology Conference on Performance Evaluation and Benchmarking* (pp. 17-32). Cham: Springer Nature Switzerland.

Raman, A., Sarkar, S., Olma, M., & Athanassoulis, M. (2023). Indexing for Near-Sorted Data. In *Proceedings of the IEEE International Conference on Data Engineering (ICDE)*.

Raman, A., Karatsenidis, K., Xie, S., Olma, M., Sarkar, S., & Athanassoulis, M. (2025, March). QuIT your B⁺-tree for the Quick Insertion Tree. In *Proceedings of the International Conference on Extending Database Technology (EDBT)*.

Acknowledgements

I would like to thank Teona Bagashvili, Anwesha Saha, Aneesh Raman, Prof. Manos Athanassoulis for their help and mentorship throughout the project. I would also like to thank Shrish Senthilkumar and Can Gokmen for their help and thoughtful discussions.

