TSM-Bench: Benchmarking Time Series Database Systems for Monitoring Applications

Abdelouahab Khelifati, Mourad Khayati, Anton Dignós, Djellel Difallah, and Philippe Cudré-Mauroux

Goal and Contribution

Motivation: Existing Time Series Database Systems (TSDBs) benchmarks are limited in the number of evaluated systems, the type of workloads, the size and type of data, and the query variability.

Goal: A comprehensive benchmark of TSDBs for monitoring applications.

Contributions:
1. Extensive evaluation of seven popular TSDBs using temporal workloads.
3. Recommendations for understanding and navigating systems’ architectural designs.

Time Series Generator

- A new generation technique that combines GAN with LSH.
- Scalable data generation of large realistic time series.

TSM Architecture

- TS-LSH uses sample data to generate large data streams.
- The executor launches configurable workload tiers.
- The statistics collection module records the performance of the TSDB.

Experiments

- The performance of the systems depends on the size of the input/output data.
- The offline and online workloads show different trends.
- Time series features heavily impact systems’ compression capability.

Performance Summary

- Seven discriminative dimensions.
- Performance ranking for different query types on a 0-5 scale.
- No silver bullet.
- Clickhouse and extremeDB offer the best trade-off.

Additional Info

- Github: https://github.com/eXascaleInfolab/TSM-Bench
- Related works:

Applications

Monitoring of Watercourse (BAFU)

- BAFU monitors the water discharge and level in Swiss rivers.
- Evaluates water quality.
- Assesses the impact of climate change and triggers alerts in case of hazard.

Other applications: Internet of Things (IoT), smart grids, traffic networks, etc.

Presented at the 49th International Conference on Very Large Databases (PVLDB), Vancouver, Canada, August 2023.