JING LAN

✉ jlan22@cse.cuhk.edu.hk · ☎ (+852) 6401-8176 · ⚙ Jing Lan

🎓 EDUCATION

The Chinese University of Hong Kong (CUHK), Hong Kong, China   2022 – Present
MPhil in Computer Science
Supervisor: Prof. James Cheng

Sun Yat-sen University (SYSU), Canton, China   2018 – 2022
BEng in Computer Science
Supervisors: Prof. Nong Xiao & Prof. Dan Huang
GPA: 3.9 / 4.0

👥 EXPERIENCES

Flavius: Disaggregated Graph OLAP Engine @ Husky Data Lab, CUHK   2022 – Present
core member   Advisor: Prof. James Cheng

We build an in-memory graph OLAP engine with a **disaggregated architecture**. The system optimizes for highly skewed graph workloads. It processes dominant small queries with high throughput while ensuring QoS for the heavy. Major Contributions:

- A *disaggregated* layout with a compute cluster connected to clients, planning and scheduling queries, and a cache cluster caching data, executing pushdown operators (e.g., filter, aggregation) and distributed joins.
- A set of novel *Subquery* operators executing optional match, pattern filter, repeated join, etc., and a query planner translating graph queries to operator pipelines.
- A *configurable*, push-pull hybrid query engine. It executes heavy queries in batches and reduces resource consumption. I co-design the query controller and the *subquery* pipeline, allowing fine-grained dataflow coordination for complex graph queries (e.g., with loop structures).
- An in-database *profiling* component collecting networking, queueing, and processing statistics when messages flow down the pipeline. It reveals query characteristics in a distributed context.
- (ongoing) A data-driven *scheduler* on-the-fly configuring our engine for performance objectives.

Student Cluster Competitions (SCC) @ SYSU   2020 – 2022
leader(2021-22), member(2020-21)   Advisor: Prof. Dan Huang

- We exploit HPC systems and win contests in world-famous venues (e.g., ACM SC & European ISC)
- Cluster management from configurable clouds (e.g., TACC Chameleon Cloud) to supercomputers (e.g., UToronto Niagara): In preparation phase, we use package managers such as Spack and Modules to manage our software stack consisting of complex, multi-version dependencies and various build systems.
- Tuning system software and parallel applications:
  - We leverage *system features*, e.g., placing threads with NUMA awareness and
  - exploit *application patterns*. At SCC@ISC, we complete an MPI profiler to capture MPI communications and visualize the I/O patterns between processes. We then balance machine I/O for the WRF Model by adjusting the process/thread layout.
- In-depth optimizations from various aspects:
  - *Parallelism* resolves performance bottlenecks by introducing multi-core and GPU processing.
  - *Code optimizations* improve temporal and spatial localities of *kernels* with tiling, vectorization, etc.

⚙ SKILLS

- Programming Languages: C, C++, Python
- Platform experience: clusters, cloud, supercomputers
- Software: Slurm, Spack, Git, CMake, tmux, LaTeX, etc.
- Languages: English (TOEFL 105, S24), Mandarin

🎓 PUBLICATIONS

TBD.

♡ Honors

2ⁿᵈ Prize (top 10%), Undergraduate Scholarship   2019, 2021
Champion   IndySCC@Supercomputing, 2021
Honorable Mention   SCC@ISC High Performance, 2021
2ⁿᵈ Prize (3ʳᵈ out of 65, and a $9000 bonus!)   Sugon Priority Research Application, 2021

🏛 Teaching

CSCI1540   Fundamental Computing With C++   Fall 2022

🛈 Miscellaneous

- The SYSU SCC Team: https://scc.sysu.tech
- Homepage: https://lan-jing.github.io/