Kushagra **Shah**

Database Systems · Computer Architecture · Embedded Systems

(+49) 1741706118 | kushagrashah298@gmail.com | kushagrashah.github.io | kushagrashah298

Summary

I aspire to contribute to the field of energy-efficient, high-performance, and data-driven computing. My interests span database systems, computer architecture, embedded systems and machine learning. I am motivated to develop cutting-edge technologies and embrace new challenges!

Experience

Huawei Technologies — *Munich, Germany*

Senior Research Engineer | Cloud Storage Team — *Jul 2023 - Present*

- Researched, prototyped and delivered a vector+scalar composite index for a distributed multi-tenant vector database in Huawei Cloud.
- Led the project, taking full responsibility for development and delivery while incorporating valuable feedback from the team and stakeholders.
- Achieved up to 5x performance improvement compared to the state-of-the-art HNSW index, while maintaining similar accuracy and index size.
- Improved the debuggability of a Spark-based system by collecting relevant statistics on run-time and updating the history server web interface.

Oracle Switzerland — *Zurich, Switzerland*

Research Assistant | Data Plane Team — *Sep 2022 - Mar 2023*

- Explored various architecture choices to optimize the data load operation in an analytical query engine while collaborating with multiple teams.
- Conducted in-depth hardware performance experiments using various benchmarks to test the viability of the proposed engine architecture.
- Experimented with storage technologies, MySQL features, data storage formats, page organization and code optimization at various levels.
- Developed a prototype which scales with data size and compute, while offering more than 3x performance improvement in the load speed.

École Polytechnique Fédérale de Lausanne (EPFL) — *Lausanne, Switzerland*

Research Scholar | Processor Architecture Lab — *Jul 2021 - Aug 2022*

- Enhanced Dynamatic (dynamic HLS) by incorporating speculation as an optimization layer, which further improved the design wrt static HLS.
- Designed custom speculative components in VHDL to enable elastic and generic speculation on any component within a dataflow graph.
- Developed an algorithm to identify speculative paths in a dataflow graph and optimize the placement of custom components for least overhead.
- Automated the process of converting a dataflow graph into a VHDL design capable of speculation, thus enabling a new feature in Dynamatic.

Selected Projects

Gradient Compression with New Numerical Encodings — *EPFL, Switzerland*

Advanced Multiprocessor Architecture Course

- Investigated the effectiveness of gradient compression on DNN models trained on a hardware emulator that uses hybrid block FP encoding.
- Tested the Python design with several image classification experiments (ResNet18 on CIFAR10) with different hyperparameters. Achieved an accuracy of about 94% with HBFP (cf. 94.7% with FP) despite using a lower precision encoding with 4 bits only.
- Implemented an RTL design for the gradient compression block which will serve as the foundation for integrating HBFP on a GPU cluster.

Distributed Execution of Near-Neighbour Algorithms — *EPFL, Switzerland*

Database Systems Course

- Implemented near-neighbour algorithm in Scala over Apache Spark, and evaluated the performance to confirm theoretical hypotheses.
- Designed a Locality Sensitive Hashing based method which uses MinHash to compute approximate results, but it is faster than the naïve version.
- Improved the algorithm – using load balancing, broadcasting, AND OR composition – and analysed the results to discover specific scenarios where each improvement gives the best results.

Skills

Industry Knowledge	System Design, Machine Learning, Containerization (Docker), Kubernetes, Digital Logic Simulation
Programming Languages	Python, C, C++, Java, Scala, Golang, Verilog, VHDL, Chisel, Shell, SQL, JS, CSS, HTML
Spoken Languages	English (C1), Hindi (C2), Gujarati (C2), French (A1), German (A1)

Education

École Polytechnique Fédérale de Lausanne (EPFL) — *Lausanne, Switzerland*

M.Sc. in Computer Science, Specialization in Data Analytics, **CGPA:** 5.3/6 — *Sep 2020 - Mar 2023*

- Thesis at Oracle, Switzerland: Revisiting Data Ingestion for a Distributed Query Engine
- Research Scholar at the Processor Architecture Laboratory (LAP)

Birla Institute of Technology and Science (BITS) Pilani — *Goa, India*

B.E. in Electrical and Electronics Engineering, **CGPA:** 9.3/10, **Rank:** 3/83 — *Aug 2016 - May 2020*

- Thesis at NTU, Singapore: Published a Technique for Vendor and Device Agnostic Hardware Area-Time Estimation
- Teaching Assistant for courses: Computer Architecture, Microprocessors and Interfacing, Digital Design