

VIRAJ THAKKAR

Ph.D. Student, Arizona State University, Tempe

viraj.online@asu.edu

• veedata.github.io

• linked.in/veedata

My research spans across the domains of Storage Systems, Data Infrastructures, Disaggregated Infrastructure, and AI/ML with Storage Systems. Feel free to reach out if something catches your eye.

RESEARCH INTERESTS

Data Infrastructure: LSM-based Key-Value Stores (RocksDB); Timeseries databases (InfluxDB); Disaggregated Memory, Storage, and Infrastructure; Storage for Quantum Computing.

Intelligent Storage: AI/ML for Storage Systems; Storage Systems for AI/ML; Real-Time Storage Systems for High-Velocity and High-Volume Data.

EDUCATION

Ph.D. Computer Science

Arizona State University

ASU Fulton Fellowship Award Recipient

08/2023 - Ongoing

Advisor: [Dr. Zhichao Cao](#)

M.S. Computer Science

Arizona State University

Thesis: Optimizing Consistency and Performance Trade-off in Distributed LSM-KVS. [URL](#)

01/2022 – 08/2023

Advisor: [Dr. Zhichao Cao](#)

3.87 / 4.00

RECENT PUBLICATIONS

3. **[SIGMOD'25] Viraj Thakkar**, Dongha Kim, Yingchun Lai, Hoekun Kim, Zhichao Cao. "SHIELD: Encrypting Persistent Data of LSM-KVS from Monolithic to Disaggregated Storage." *Proceedings of ACM Conference on Management of Data (SIGMOD), Research Track Full Paper, 2025. To Appear.*
2. **[HotStorage '24] Viraj Thakkar**, Madhumitha Sukumar, Jiaxin Dai, Kaushiki Singh, and Zhichao Cao. 2024. "Can Modern LLMs Tune and Configure LSM-based Key-Value Stores?" *In Proceedings of the 16th ACM Workshop on Hot Topics in Storage and File Systems, July 08, 2024. ACM, Santa Clara CA USA, 116–123. [URL](#).*
BEST PAPER AWARD
1. **[SIGMOD '24] Qiaolin Yu**, Chang Guo, Jay Zhuang, **Viraj Thakkar**, Jianguo Wang, Zhichao Cao. "CaaS-LSM: Compaction-as-a-Service for LSM-based Key-Value Stores in Storage Disaggregated Infrastructure." *Proceedings of ACM Conference on Management of Data (SIGMOD), Research Track Full Paper, 2024. [URL](#).*

ACADEMIC POSTERS AND WORK-IN-PROGRESS

2. **[FAST '25] Yibo Zhao**, **Viraj Thakkar**, Zhichao Cao, Zaoxing Liu. "NetLSM: Enabling an In-Network Approach for Scheduling LSM-KVS Operations." 23rd USENIX Conference on File and Storage Technologies, 2025.
1. **[FAST '24] Madhumitha Sukumar**, Jiaxin Dai, Kaushiki Singh, Vikriti Lokegaonkar, **Viraj Thakkar**, Zhichao Cao. "LLM-assisted Automatic-Configuration and Tuning Framework for LSM-based Key-Value Stores." 22nd USENIX Conference on File and Storage Technologies, 2024

EXPERIENCE

System Software Research Associate Intern

HP Labs

- Ongoing :)

05/2025 – Ongoing

California., USA

Software Development Engineer Intern

Amazon

- Set up product documentation to seamlessly onboard future team members.
- Worked with the team to develop a full-stack application (AWS Lambda, DynamoDB, S3 buckets, ECS, SQS CloudFormation) that provides an interface to check for copyright infringements for products on Amazon.

05/2023 – 08/2023

Washington D.C., USA

Junior Researcher

Somaiya Vidyavihar, IN & Children's Investment Fund Foundation (CIFF), U.K.

- Led team of six to explore change brought by technology towards Online Child Sexual Abuse (OCSA) and Exploitation (OCSE), presenting a comprehensive work to leading voices across the country.
- Developed a tool to automate search and filter instances of OCSA and OCSE in Legal Cases.

09/2019 - 03/2020

Mumbai, IN

TECHNICAL SKILLS

Languages:	Python, C/C++, Shell, SQL, MATLAB, Front End Development (HTML, CSS, JS)
Databases:	LSM-KVS (RocksDB, LevelDB), Graph DB (Neo4j), Timeseries (InfluxDB), Misc. (MySQL, SQLite, PostgreSQL, HDFS)
Tools & More:	Git, Kubernetes, Docker, Django, AWS, Azure, IBM Quantum
Soft Skills:	Project Management, Leadership, Consistency, Communication Skills, Creative problem-solving

PROJECTS IN PROGRESS

Automating LLM-Assisted Tuning for LSM-KVS

- LSM-KVS (like RocksDB) have 100+ interrelated options that need deep understanding to be tuned correctly.
- Researched and developed a framework that automates predicting workloads and setting up tuning parameters for LSM-KVS by leveraging corpus ingestion and inferencing abilities of Modern LLMs.

Leveraging Programmable Network Switch for pre-emptive LSM-KVS Operations

- LSM-KVS exposes options that can be dynamically modified at runtime to improve performance.
- Explored and developed a controller to enhance LSM-KVS performance using programmable network switches (P4 Switches) to perform operations and online configuration modifications for LSM-KVS.

An Exploration of High Throughput Timeseries Databases for Cutting-Edge Devices

- Timeseries databases require high write-ingestion rates to handle real-time data streams. This research is currently in the early stages of exploring the integration of ZNS devices for timeseries databases like InfluxDB.

Integrating Encryption into LSM-KVS in Monolith and Disaggregated Infrastructure

- Encryption of persistent data in LSM-KVS provides data confidentiality and is essential as deployments move from monolith to multi-tenant disaggregated infrastructures.
- Designed and developed scalable and decentralized encryption design for LSM-KVS that works from monolith to disaggregated infrastructure while embedding practices like encryption key rotation into LSM-KVS components.

GRANTS AND AWARDS

ACM SIGMOD 2025 Travel Grant	2025
ASU SCAI, Graduate College, and GSG Travel Grant	2025
OpenAI Researcher Access Program Credit Award	2025
USENIX FAST 2024, 2025 Travel Grant	2024, 2025
ASU Fulton Fellowship	2023 - 2027

ACADEMIC SERVICES

Shadow Program Committee Member,	<i>EuroSys 2025</i>
Artifact Evaluation Committee,	<i>FAST 2025</i>
Artifact Evaluation Committee,	<i>SIGMOD 2025</i>

SOCIAL CONTRIBUTIONS AND ACHIEVEMENTS

CGroup-Manager , A Python library to start and monitor a Linux process using cgroups. URL	2024
Stack Overflow , URL	2023
• 1200+ Reputation with 50+ answers. <i>Top 15% contributor in 2023.</i>	
Hackathon by SVB & ASU	2022
• Developed a full stack web application that connects Investors and Startups. <i>Runner-Up.</i>	
Memories , A Python library for Image Segmentation. URL	2022
Internship and Hackathon Organizer , Mumbai, IN	2020 - 2021
• Designed and organized a college-wide internship for 100+ interns over 2 semesters.	
• Worked with the college's computer science society to manage an annual College Hackathon.	
Qubit by Qubit , a course on Quantum Computing by IBM Quantum and The Coding School	2021
Team KJSCE Robocon , Mumbai, IN	2018 - 2020
• Spearheaded a team of 40+ individuals and mentored members to take over the team in coming years.	
• Achieved National Rank 8 and Winner of Best Solution Idea amongst 50+ teams in "DD-Robocon 2020."	