

# VIRAJ THAKKAR

Ph.D. Student, Arizona State University

[viraj.dt@asu.edu](mailto:viraj.dt@asu.edu)

• [veedata.github.io](https://github.com/veedata)

• [linkedin.com/veedata](https://linkedin.com/veedata)

*My research focuses on the intersection of Storage Systems in Modern Disaggregated Infrastructure Deployments.*

*Feel free to reach out if something catches your eye :)*

## RESEARCH INTERESTS

---

**Data Infrastructure:** LSM-based Key-Value Stores (RocksDB, LevelDB); Disaggregated Compute, Memory, and Storage; Storage for Quantum Computing; Data Confidentiality.

**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*

Aug. 2023 - Ongoing

Advisor: [Dr. Zhichao Cao](#)

3.95 / 4.00

### M.S. Computer Science

Arizona State University

*Thesis: Refining Consistency and Performance Trade-off in Distributed LSM-KVS. [\[PDF\]](#)*

Jan. 2022 – Aug. 2023

Advisor: [Dr. Zhichao Cao](#)

3.87 / 4.00

### B.Tech. Information Technology

K. J. Somaiya College of Engineering, India

Jan. 2017 – Sept. 2021

8.56 / 10.00

## HONORS, GRANTS, AND AWARDS

---

ACM SIGMOD Travel Grant

2025

ASU SCAI, Graduate College, and GSG Travel Grant

2025, 2026

OpenAI Researcher Access Program Credit Award

2025

HotStorage 2024 SK Hynix Best Paper Award

2024

USENIX FAST Travel Grant

2024, 2025, 2026

ASU Fulton Fellowship

2023 - 2027

## PEER-REVIEWED PUBLICATIONS

---

11. **[SeQureDB '26] Viraj Thakkar**, Dongha Kim, Hokeun Kim, Zhichao Cao. "ANCHOR: A Vision for Secure Persistent Key-Value Stores in Disaggregated Data Centers." *1st SeQureDB Workshop on Secure & Private Data Management, 2026. Co-located with ACM SIGMOD 2026. [To Appear.](#)*
10. **[SIGMOD '26]** Qi Lin, Gangqi Huang, Te Guo, Chang Guo, **Viraj Thakkar**, Zhichen Zhu, Jianguo Wang, Zhichao Cao. "O3-LSM: Maximizing Disaggregated LSM Write Performance via Three-Layer Offloading." *Proceedings of ACM Conference on Management of Data (SIGMOD), Research Track Full Paper, 2026. [To Appear.](#)*
9. **[HPE TechCon '26]** John Bryne, **Viraj Thakkar**, Matthias Frank, Elias Klett, Harumi Kuno, Sekwon Lee, Clarete Crasta, Abhishek Dwaraki, David Emberson, Ramya Rao, Amitha C, Chinmay Ghosh, Prono Shome, Shreyas Basri. "CacheFS." *Published at HPE internal conference. 82 of 1036 (8%) Acceptance Rate. [Close Access.](#)*
8. **[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. [\[PDF\]](#)*
7. **[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 16<sup>th</sup> ACM Workshop on Hot Topics in Storage and File Systems, July 08, 2024. ACM, Santa Clara CA USA, 116–123. [\[PDF\]](#)*  
**BEST PAPER AWARD!**

6. [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*. [\[PDF\]](#)
5. [RPA '23] Ashwini Dalvi, **Viraj Thakkar**, Smit Moradiya, Aditya Vedpathak, Irfan Siddavatam, Faruk Kazi, S. G. Bhirud. "FLASH: Web-Form's Logical Analysis & Session Handling Automatic Form Classification and Filling on Surface and Dark Web." *Robotic Process Automation (2023): 61-100*. [\[PDF\]](#)
4. [HRAI-HCI '22] Irfan Siddavatam, Ashwini Dalvi, Abhishek Patel, Aditya Panchal, Aditya S. Vedpathak, **Viraj Thakkar**. "ML-Enabled Informed Intervention for Crowdsourcing-Based Optimization of Medical Resources." *Handbook of Research on Applied Intelligence for Health and Clinical Informatics*. IGI Global, 2022. [\[PDF\]](#)
3. [ICACIT '21] Ashwini Dalvi, Irfan Siddavatam, **Viraj Thakkar**, Aditya Vedpathak, Abhishek Patel. "Recognizing Child Unsafe Apps Through User Reviews on the Google Play Store." *Advanced Computing and Intelligent Technologies* (pp. 111-120). Springer, Singapore 2022. [\[PDF\]](#)
2. [ICAST '21] Irfan Siddavatam, Ashwini Dalvi, **Viraj Thakkar**, Aditya Vedpathak, Smit Moradiya, Apoorva Jain. "Explainability using decision trees and Monte Carlo simulations." *Proceedings of the 4th International Conference on Advances in Science & Technology (ICAST2021)*. [\[PDF\]](#)
1. [IBSSC '21] Ashwini Dalvi, Irfan Siddavatam, **Viraj Thakkar**, Apoorva Jain, Faruk Kazi, Sunil Bhirud. "Link Harvesting on the Dark Web." *IEEE Bombay Section Signature Conference (IBSSC)*, pp. 1-5. IEEE, 2021. [\[PDF\]](#)

### PEER-REVIEWED POSTERS AND ABSTRACTS

---

5. [HPTS '26] **Viraj Thakkar**. "Broad Reasoning but Narrow Control for an End-to-End KVS Tuner." *21st International Workshop on High Performance Transaction Systems (HPTS)*. [Under Review](#).
4. [FAST '26] Zhenjie Sun, **Viraj Thakkar**, Avani Wildani, Zhichao Cao. "Beyond ARC – Introducing LLM-driven Adaptive Policy Replacement." *24th USENIX Conference on File and Storage Technologies, 2026*.
3. [CUG '26] John Bryne, Myungjun Son, Harumi Kuno, Aditya Dhakal, Sekwon Lee, **Viraj Thakkar**, Clarete Crasta, David Emberson. "Characterizing sustained low-level message latency." *Conference for the Cray User Group, 2026*.
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*.

### PREPRINTS / UNDER REVIEW

---

4. [VLDB '27] **Viraj Thakkar**, Qi Lin, Kenanya Prasetyo, Raden Wisjnunandono, Jin Wang, Achmad Kistijantoro, Reza Rachmadi, Zhichao Cao. "TellyTune: Exploring the Limits of LLM-Driven Tuning for LSM-Based Key-Value Stores." *53rd International Conference on Very Large Data Bases, 2027*. [Under Review](#).
3. [HPTS '26] **Viraj Thakkar**. "Broad Reasoning but Narrow Control for an End-to-End KVS Tuner." *21st International Workshop on High Performance Transaction Systems (HPTS)*. 2026. [Under Review](#).
2. [arXiv '25] **Viraj Thakkar**, Qi Lin, Kenanya Keandra Adriel Prasetyo, Raden Haryosatyo Wisjnunandono, Achmad Imam Kistijantoro, Reza Fuad Rachmadi, Zhichao Cao. "ELMo-Tune-V2: LLM-Assisted Full-Cycle Auto-Tuning to Optimize LSM-Based Key-Value Stores." *arXiv:2502.17606*. 2025. [\[PDF\]](#)
1. [arXiv '25] Qi Lin, Zhenyu Zhang, **Viraj Thakkar**, Zhenjie Sun, Mai Zheng, Zhichao Cao. "StorageXTuner: An LLM Agent-Driven Automatic Tuning Framework for Heterogeneous Storage Systems." *arXiv:2510.25017*. 2025. [\[PDF\]](#)

### PATENTS AND INVENTION DISCLOSURES

---

2. John Bryne, **Viraj Thakkar**, Harumi Kuno, Sekwon Lee. "A Global, Strongly Consistent, Client-Sided Caching Mechanism." [Under HPE Internal Disclosure Review, Patent filing upon Approval](#).
1. John Bryne, **Viraj Thakkar**, Harumi Kuno, Sekwon Lee. "Federated Coherence Management for a Globally Shared Cache." [Under HPE Internal Disclosure Review, Patent filing upon Approval](#).

## INVITED TALKS

---

6. "ANCHOR: A Vision for Secure Persistent Key-Value Stores in Disaggregated Data Centers." *1st SeQureDB Workshop on Secure & Private Data Management, 2026. Upcoming.*
5. "SHIELD: Encrypting Persistent Data of LSM-KVS from Monolithic to Disaggregated Storage." *ACM Conference on Management of Data (SIGMOD), 2025.*
4. "NetLSM: Enabling an In-Network Approach for Scheduling LSM-KVS Operations." *23rd USENIX Conference on File and Storage Technologies, 2025.*
3. "Can Modern LLMs Tune and Configure LSM-based Key-Value Stores?" *16th ACM Workshop on Hot Topics in Storage and File Systems HotStorage'24, CA, 2024.*
2. "LLM-assisted Automatic-Configuration and Tuning Framework for LSM-based Key-Value Stores." *22nd USENIX Conference on File and Storage Technologies, 2024.*
1. "Explainability using decision trees and Monte Carlo simulations." *4th International Conference on Advances in Science & Technology, 2021.*

## ACADEMIC SERVICES

---

### Shadow Program Committee,

- SIGMETRICS 2026
- EuroSys 2025

### Artifact Evaluation Committee,

- FAST 2026
- FAST 2025
- SIGMOD 2025

## MENTORSHIP

---

<b>2026</b>	<i>MS Student</i>	Sampada Nemade ( <i>Modeling Traces using Statistical Methods</i> )
<b>2025</b>	<i>HPE Labs Dual-Study Intern</i>	Elias Klett ( <i>Benchmarking FUSE as an In-Memory Filesystem</i> )
	<i>MS Student</i>	Anushwar Sharma ( <i>The Fusion of LLMs and ML for LSM-KVS tuning</i> )
<b>2024</b>	<i>MS Thesis Student</i>	Harry Deccoco ( <i>Optimizing Post-deletion performance in LSM-KVS</i> )
<b>2023</b>	<i>Undergrad Project Group</i>	Madhumita Sukumar; Kaushiki Singh ( <i>LLM-driven tuning for LSM-KVS</i> )
	<i>MS Project Group</i>	Jiaxin Dai ( <i>LLM-driven tuning for LSM-KVS</i> )
<b>2020</b>	<i>Undergrad Robotics Club</i>	Anish Pawar; Kamal Rohra; Dhruv Joshi ( <i>Simulating object physics in MATLAB</i> )

## EXPERIENCE

---

### HPE Labs, Systems Architecture Lab

System Software Research Associate Intern

*California, USA*

May 2025 – Ongoing

- FUSE through the Looking Glass, A Modern Take.
- A client-sided global distributed cache!
- More ongoing timey-wimey fun things.

### Amazon

Software Development Engineer Intern

*Washington D.C., USA*

May 2023 – Aug. 2023

- Established product documentation to seamlessly onboard future team members.
- Collaborated 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.

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

Junior Researcher

*Mumbai, IN*

Sept. 2019 – Mar. 2020

- Led team of six to develop an automated tool identifying Legal Cases about Online Child Sexual Abuse (OCSA) and Exploitation (OCSE) in India, presenting a comprehensive work to leading voices across the country.
- Collaborated with the Indian National Commission for Protection of Child Rights on a report that was submitted to the Rajya Sabha (Upper House) of the Indian Parliament.

## PROJECTS IN PROGRESS

---

### Automating LLM-Assisted Tuning for LSM-KVS

- LSM-KVS (like RocksDB) have 200+ 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.

### 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.

## SOCIAL CONTRIBUTIONS AND ACHIEVEMENTS

---

<b>CGroup-Manager</b> , A Python library to start and monitor a Linux process using cgroups. <a href="#">[URL]</a>	2024
<b>Stack Overflow</b> , <a href="#">[URL]</a>	2023
<ul style="list-style-type: none"><li>• 1200+ Reputation with 50+ answers. <i>Top 15% contributor in 2023.</i></li></ul>	
<b>Hackathon by SVB &amp; ASU</b>	2022
<ul style="list-style-type: none"><li>• Developed a full stack web application that connects Investors and Startups. <i>Runner-Up Prize.</i></li></ul>	
<b>Memories</b> , A Python library for Image Segmentation. <a href="#">[URL]</a>	2022
<b>Internship and Hackathon Organizer</b> , Mumbai, IN	2020 - 2021
<ul style="list-style-type: none"><li>• Designed and organized a college-wide online internship for 100+ interns over 2 semesters.</li><li>• Worked with the college's computer science society to manage an annual College Hackathon.</li></ul>	
<b>Qubit by Qubit</b> , a course on Quantum Computing by IBM Quantum and The Coding School	2021
<b>CoView, Interactive COVID-19 Portal for KEM Hospital</b> , Mumbai, IN	2020
<ul style="list-style-type: none"><li>• Implemented interactive dashboard for KEM Hospital to monitor patients in the COVID-19 pandemic.</li><li>• CoView was deployed and used for over 7 months during the pandemic. Featured in local news <a href="#">[URL]</a>.</li></ul>	
<b>Team KJSCE Robocon</b> , Mumbai, IN	2018 - 2020
<ul style="list-style-type: none"><li>• Spearheaded a team of 40+ individuals and mentored members to take over the team in coming years.</li><li>• Achieved National Rank 8 and Winner of Best Solution Idea amongst 50+ teams in "DD-Robocon 2020."</li></ul>	