

# Omkar Desai

[omkarbdesai@gmail.com](mailto:omkarbdesai@gmail.com) | +1 (201) 442-9613 | [linkedin.com/in/omkar-desai/](https://www.linkedin.com/in/omkar-desai/) | [github.com/swiftomkar](https://github.com/swiftomkar)

Systems researcher specializing in cost-efficient, fault-tolerant distributed ML training. Published at USENIX FAST, EuroMLSys, HPDC, and HotStorage. Two granted US patents. Open-source contributor to PyTorch and Linux BPF.

## EDUCATION

Ph.D. in Computer Science (Systems Research), *Syracuse University* Expected May 2026  
Master of Engineering, *Stevens Institute of Technology* 2019  
Bachelor of Engineering, *University of Mumbai* 2017

## EXPERIENCE

**Research Assistant** - *Syracuse University, Syracuse, NY* Aug 2020 – Present  
*Systems research lab – Advised by: Bryan Kim*

- **EuroMLSys 2026 publication:** Built a spot-instance orchestration system for distributed LLM training that integrates adaptive checkpointing and preemption-aware scheduling, reducing training cost by 2.04x
- **FAST 2026 publication:** Designed a caching and data preprocessing system for ML training and improved performance by 3.45x
- **HPDC 2023 publication:** Optimized Key-Value SSDs for production workloads using RocksDB
- **HotStorage 2022 publication:** Characterized block I/O workloads and predicted I/O performance of a device
- Optimizing PyTorch and MegatronLM for ML training by improving the performance of data movements between the host and GPUs
- Paper reviewer for ACM Transactions on Storage; Artifact evaluator for USENIX FAST '25 and NSDI '26

**PhD Research Intern** - *Samsung Semiconductor, Inc., San Jose, CA* May 2022 – Aug 2022  
*Memory Solutions Lab*

- Built an all-flash array simulator for NVMe SSDs (MQSim+QEMU) supporting up to 256-drive RAID configurations, enabling accurate application I/O simulation
- **US patent 12411630:** Designed tools to identify and mitigate long tail latencies in AFA storage systems. Patented an AI model to predict GC and divert I/Os away, reducing tail latencies by a magnitude

**PhD Research Intern** - *Samsung Semiconductor, Inc., San Jose, CA* May 2021 – Aug 2021  
*Memory Solutions Lab*

- Researched near-data processing using computational storage devices; proposed a memory-efficient learned index for key-value SSDs
- **US patent 11954345:** Integrated learned indexing with traditional SSD indexing to reduce memory footprint while minimizing the compute overhead of retraining ML models

**Software Engineer** - *Practo Technologies, Bangalore, India* Jun 2019 – Sep 2020  
*DevOps group*

- Designed the Kubernetes and data engineering infrastructure for the largest healthcare platform in India with 20 million daily active users
- Developed a scalable user behavior and events pipeline to ingest 100K events/min from the website to customize recommendations for users
- Led development of ETL tools for data pipelines and used them for pre-processing data for training DLRM models

**Software Engineering Intern** - *Delos Living, New York, NY* May 2018 – Dec 2018  
*Delos Labs*

- Collaborated with clinicians to build an IoT-powered indoor wellness ecosystem using smart sensors and appliances for R&D
- Designed and developed a scalable cloud-native platform on AWS, capable of supporting up to 10,000 homes to facilitate large-scale testing and evaluation efforts at the R&D lab.

## PUBLICATIONS

- **Omkar Desai**, S. Pei, J. Bhimani, B. Kim. Cost-Efficient Training and Checkpointing for Large Models on Preemptible Cloud VMs (EuroMLSys 2026)
- **Omkar Desai**, Z. Jiao, S. Pei, J. Bhimani, and B. Kim. 2025 Preparation Meets Opportunity: Enhancing Data Preprocessing for ML Training With Seneca (USENIX FAST 2026) <https://www.usenix.org/conference/fast26/presentation/desai>
- Z. Jiao, **Omkar Desai**, J. Kim, J. Choi, B. Kim. paRAID: An Efficient and Sustainable Storage Architecture with Heterogeneous SSDs [Under review]
- M. Saha, **Omkar Desai**, B. Kim, and J. Bhimani. 2023. Leveraging Keys In Key-Value SSD for Production Workloads. (HPDC '23) <https://dl.acm.org/doi/abs/10.1145/3588195.3595949>
- **Omkar Desai**, S. Shin, E. Lee, and B. Kim. 2022. A principled approach for selecting block I/O traces. (HotStorage '22) <https://doi.org/10.1145/3538643.3539754> (Won best presentation award)

- **Omkar Desai** and Parth Bhatt. 2017. Home Automation Using Wi-Fi and Sensor Networks. *Technofocus*, ISSN:2321-0532

## **PATENTS**

---

- **Omkar Desai**, Z. Jiao, S. Pei, J. Bhimani, and B. Kim. Preparation Meets Opportunity: Enhancing Data Preprocessing for ML Training [US patent pending]
- **Omkar Desai**, C. Choi, and Y. Kang. Two level indexing for key-value persistent storage device [US patent: 11954345]
- **Omkar Desai**, S. Pei, J. Yang, and R. Pitchumani. System and Method for Managing Tasks in a Storage Device [US Patent: 12411630]

## **AWARDS, and SERVICE**

---

- ACM Best presentation award – HotStorage 2022
- Best poster award – Syracuse University ECS research day 2024
- Paper reviewer for ACM Transactions on Storage in 2026
- Artifact evaluator USENIX FAST 2025
- Artifact evaluator USENIX NSDI 2026

## **SKILLS**

---

- **Domains:** Distributed Systems, Computer architecture, SSD firmware, GPUs, AI/ML systems, HPC, Databases, Storage systems, Operating Systems, Data engineering, Flash storage
- **Open-Source software contributions:** PyTorch, MegatronLM, Apache Airflow, Linux strace, Berkeley Packet Filter (BPF)
- **Systems:** File Systems, RAID, Ceph, Lustre, ZFS, LevelDB, RocksDB, Kubernetes, Apache Airflow, Kafka, Hadoop
- **Languages & Frameworks:** C++, C, Python, PyTorch, CUDA, TensorFlow, Flask, SQLAlchemy, Java, Spring, Redis
- **Tools:** Berkeley Packet Filter, Blktrace, strace, AWS, Git, Docker, etc.

## **TEACHING & MENTORING**

---

**Syracuse University**, Teaching Assistant

Aug 2023 – May 2024

- Lead teaching assistant for undergraduate and graduate Operating Systems class (CIS 657 and CIS 486)

**Syracuse University**, Graduate Mentor

May 2023 – Present

Closely mentored BS and MS summer research interns on a research project focused on training systems for transformer-based LLMs