Sandeep **Kumar**

RESEARCH SCIENTIST @ INTEL, BENGALURU, INDIA

--

📱(+91) 8277361995 | ✉ sandeep007734@gmail.com | 🏠 https://sandeep007734.github.io | 📮 sandeep007734 | in sandeep007734

Professional Summary

Systems researcher specializing in **AI/ML performance optimization** and **memory tiering**. Expertise in memory management for DNN training (PyTorch), including Transformer and CNN architectures, LLM inference, and RAG systems. Proven track record of optimizing memory-intensive workloads through tiered memory architectures, kernel-level enhancements, and hands-on experience with hardware accelerators (Intel IAA, DSA).

Contributor to the **Linux kernel** (memory management subsystem) with upstream patches. Currently leading a research team at Intel focused on AI infrastructure optimization, CXL memory technologies, and next-generation memory tiering solutions.

Selected Publications

1. *TierScape: Harnessing Multiple Compressed Tiers to Tame Server Memory TCO*
   Sandeep Kumar, Aravinda Prasad, and Sreenivas Subramoney. In **EuroSys** , 2026.

2. *TierTrain: Proactive Memory Tiering for CPU-Based DNN Training*
   Sathvik Swaminathan, **Sandeep Kumar**, Aravinda Prasad, and Sreenivas Subramoney. In **ISMM**, Seoul, South Korea, 2025.

3. *A Tug-of-War between Static and Dynamic Memory in Intel SGX*
   Sandeep Kumar, Abhisek Panda, Advait Nerlikar and Smruti R. Sarangi. In **VLSID, Bangalore, India** , 2024.

4. *Telescope: Telemetry for Gargantuan Memory Footprint Applications*
   Alan Nair, **Sandeep Kumar**, Aravinda Prasad, Ying Huang, Andy Rudoff, Sreenivas Subramoney. In **USENIX, ATC, USA**, 2024.

5. *Enabling Migration of Large SGX Enclaves in a Data Center*
   Sandeep Kumar, Abhisek Panda, Smruti R. Sarangi. In **arXiv** , 2024.

6. *Perspector: Benchmarking Benchmark Suites*
   Sandeep Kumar, Abhisek Panda, Smruti R. Sarangi. In **DATE**, Antwerp, Belgium, 2023.

7. *SecureLease: Maintaining Execution Control in The Wild using Intel SGX*
   Sandeep Kumar, Abhisek Panda, and Smruti R. Sarangi. In **Middleware**, Quebec City, Canada, 2022.

8. *SGXGauge: A Comprehensive Benchmark Suite for Intel SGX*
   Sandeep Kumar, Abhisek Panda, and Smruti R. Sarangi. In **ISPASS**, Singapore, 2022.

9. *SecureFS: A Secure File System for Intel SGX*
   Sandeep Kumar and Smruti R. Sarangi. In **RAID**, Spain, 2021.

10. *Page Table Management for Heterogeneous Memory Systems*
    Sandeep Kumar, Aravinda Prasad, Smruti R. Sarangi, and Sreenivas Subramoney. In **ISMM**, Canada, 2021.

11. *F-LaaS: A Control-Flow-Attack Immune License-as-a-Service Model*
    Sandeep Kumar, Diksha Moolchandani, Takatsugu Ono, and Smruti Sarangi. In **IEEE SCC**, Milan, Italy, 2019.

12. *Scalable Performance Tuning of Hadoop MapReduce: A Noisy Gradient Approach*
    Sandeep Kumar, Sindhu Padakandla, Chandrashekar L, Priyank Parihar, Gopinath K, and Shalabh Bhatnagar. In **IEEE Cloud**, Hawaii, USA, 2017.

Complete list here: https://sandeep007734.github.io/publications/

Patents

1. Context-aware memory tiering for machine learning training
   Sathvik Swaminathan, Sandeep Kumar, Aravinda Prasad, and Sreenivas Subramoney

2. Methods and apparatus to profile page tables for memory management.
   Aravinda Prasad, Sandeep Kumar, Sreenivas Subramoney, and Andy Rudoff

Education

Indian Institute of Technology Delhi   *New Delhi, India*

Doctor of Philosophy   *Jul. 2017 - Jun 2024*

- Advised by Prof. S. R. Sarangi

Indian Institute of Science   *Bangalore, India*

Master of Engineering in Computer Science   *Jul. 2011 - Aug. 2013*

- Advised by Prof. K. Gopinath

Guru Gobind Singh Indraprastha University   *New Delhi, India*

Bachelor of Technology in Computer Science   *Jul. 2007 - Aug. 2011*

Work Experience

Intel Labs   *Bangalore, Karnataka*

Research Scientist   *March 2022 – Current*

- Performance optimization for AI/ML training and inference workloads; memory tiering for LLMs and RAG systems
- Linux kernel: zswap multi-tier support (TierScape), DAMON telemetry extensions (Telescope); Author: Sreenivas Subramoney
- Hardware accelerator integration (Intel IAA, DSA) for memory compression and data movement

Intel Labs   *Bangalore, Karnataka*

Graduate Research Intern   *July 2020- Jan 2021*

Indian Institute of Science   *Bangalore, Karnataka*

Research Associate   *Sept 2014- July 2017*

Dell R&D   *Bangalore, India*

Software Development Engineer   *Jul 2013-Jun 2014*

Interests & Activities

Reading   Goodreads Profile

https://goodreads.com/sandeep007734

Running, Cycling, Hiking   Strava Profile

https://www.strava.com/athletes/3569697

References

Prof. Smruti R. Sarangi   *IIT Delhi, India*

Professor, Department of Computer Science

srsarangi@cse.iitd.ac.in

Prof. K. Gopinath   *IISc Bangalore, India*

Professor, Computer Science and Automation

gopi@iisc.ac.in