Shubham Kaushik

shubhamkaushik.com | LinkedIn | GitHub | kaushiks[at]brandeis[dot]edu

EDUCATION

Jan 2024 - Present	**Doctor of Philosophy** in *Computer Science* (*GPA: 3.97/4.0*) Brandeis University, MA, United States
Sep 2022 - Dec 2023	**Master of Science** in *Computer Science "Data-Centric Computing"* (GPA: 3.88/4.0) Boston University, MA, United States
Jul 2014 - Jun 2018	**Bachelor of Technology** in *Computer Science & Engineering* Maharshi Dayanand University, India

EXPERIENCE

Jan 2024 - Present	**Ph.D. Researcher** @ Smart & Scalable Data Systems Lab, Brandeis University
Mar 2022 - Aug 2022	**Software Engineer** @ Server Programming Team, Kwalee, India
Jun 2021 - Mar 2022	**Engineer - Information Security** @ Cyber Fusion, FIS Global, India
Jul 2018 - Jun 2021	**Project Engineer** @ Cloud Computing - Wipro Digital, Wipro Limited, India

TECHNICAL SKILLS

- **Programming Languages**: C, C++, Python, SQL, Rust
- **Markup Languages**: HTML, CSS, JSON, YAML, LaTeX, Markdown
- **Databases**: RocksDB, Postgres, MongoDB, Redis, SQLite, ORM
- **Tools & Systems**: Django, Django-REST, Kafka, Hadoop, gRPC, Asyncio, Git, ETL, Flink, AWS

KEY PROJECTS

Tectonic [Paper PDF] [GitHub] **Apr 2025 - Present**

- Designed Tectonic, a Rust-based workoad generator modeling dynamic properties of real-world workloads.
- Achieved 2× higher throughput and 84% lower memory footprint compared to state-of-the-art benchmarks.

LSM Buffer Analysis [Paper PDF] [Blog] [GitHub] **Jan 2024 - Present**

- Developed and benchmarked multiple in-memory buffer designs (skiplist, vector, hash-hybrids) for Log Structured Merge Tree (LSM) based storage engines, optimizing throughput, latency, and memory usage.
- Created a framework for optimal buffer selection for various workloads, enabling data-driven design choices.

RangeReduce [Blog] [GitHub] **Sep 2023 - Present**

- Designed RangeReduce, a range query-optimized LSM engine that leverages range queries output to derive compactions, improving system performance by reducing disk reads/writes and saving CPU-cycles.
- Achieved up to 18% lower range-query latency, 90% reduction in compaction debt, 12% less overall data movement, and 20% lower space utilization, enhancing overall LSM engine efficiency.

Edge Aware Stream Processing [Blog] [GitHub] **Jan 2023 - May 2023**

- Developed a dynamic operator placement algorithm in Apache Flink to offload tasks to Raspberry Pi edge devices, leveraging performance metrics to reduce WAN data traffic.
- Prototype preprocessed data streams at the edge, improving system efficiency and demonstrating lower latency and server-side resource usage compared to conventional homogeneous deployments.

Automated Vulnerability Detection [Blog] [Report] [GitHub] **Jan 2023 - May 2023**

- Developed an automated tool to detect security risks in third-party VS Code extensions, targeting Path Traversal and Zip Slip attacks, and identified vulnerabilities in 5% of extensions analyzed.

SELECTED PUBLICATIONS / POSTERS

DBTest 2024	**Shubham Kaushik** and Subhadeep Sarkar. *Anatomy of the LSM Memory Buffer: Insights & Implications*, In Proceedings of the International Workshop on Testing Database Systems
JCSE 2019	**Shubham Kaushik** and Ratneshwer. *Fault Modeling of an Object-Oriented System using CPN*, International Journal of Computer Sciences and Engineering
NEDB Day 2025/2024	**Shubham Kaushik**, Manos Athanassoulis, and Subhadeep Sarkar *RangeReduce: A Range Query Driven Compaction for LSM-Trees*, North East Database Day (*poster*)

CURRICULAR ACTIVITIES

- Member, VLDB 2026 Shadow PC
- Judged and mentored at *HackMIT 2023* and *BostonHacks*