Kalit Inani

kinani3@gatech.edu | linkedin.com/in/kalit-inani | github.com/Kalit31 | kalit31.github.io/

Education

Georgia Institute of Technology, Atlanta Aug 2024 – May 2026
Master of Science in Computer Science (Specialization in Computing Systems) GPA: 4.0/4.0
Relevant Coursework: Distributed Computing, Advanced OS, Systems for AI, Computer Networks, IoT

Birla Institute of Technology and Science (BITS), Pilani Aug 2018 – July 2022
Bachelor of Engineering (B.E.) in Computer Science **CGPA: 9.44/10.0 (Distinction)**
- Recipient of **Institute Merit Scholarship** for being in the top 2% of 1100 students for 6 semesters.

Work Experience

Bedrock Data *Member of Technical Staff, Intern* May 2025 – Aug 2025
- Worked in a startup environment (<15 members), taking ownership of key features and managing various ad-hoc tasks.
- Architected a low-latency semantic search system over Neo4j vector DBs using HNSW, serving 50M+ graph nodes.
- Built a scalable embedding generation pipeline for real-time vectorization of data, enabling context-aware discovery.
- Optimized incremental ingestion workflows, cutting data-processing latency by 40% and improving throughput.
- Developed a comprehensive search evaluation framework, driving 85% correlation with LLM-assessed relevance.

Nutanix *Member of Technical Staff* July 2022 – July 2024; Jan 2022 – June 2022 (Intern)
- Developed orchestration workflows for PostgreSQL and MongoDB lifecycles—delivering zero-downtime maintenance and point-in-time restores—powering Nutanix Database Service (DBaaS) across on-premise environments.
- Engineered one-click provisioning for MongoDB shards, integrating OM, cutting deployment from days to under 40 mins.
- Implemented TLS encryption for high-availability clusters securing data in transit and automating certificate management
- Designed storage-scaling for MongoDB, automating disk addition and filesystem resizing to seamlessly expand capacity.
- Contributed to team-wide code refactoring and test coverage improvements, reducing customer incidents by 45%.
- Diagnosed and resolved 50+ critical bugs, delivered production features, and provided assistance in customer on-calls.
- Participated in knowledge-sharing sessions on distributed and database systems across cross-functional teams.

JP Morgan Chase & Co. *Quantitative Research Intern* May 2021 – July 2021
- Optimized C++ cash-flow generation modules by profiling function-call overheads, and reducing execution time by 30%.
- Developed regression models for bond-pricing inference using quantitative analysis and backtesting on historical data.

Indian Institute of Remote Sensing (IIRS), ISRO *Research Intern* May 2020 – July 2020
- Developed a U-Net CNN architecture for 8-class satellite image segmentation using Sentinel-2 data from Google Earth Engine, trained across 20,000+ processed images, achieving 80% validation accuracy for cyclone damage analysis.

Research and Academic Projects

- **Distributed Systems**: Built a distributed sharded key-value store using Multi-Paxos consensus for fault tolerance and 2PC for cross-shard transactions, featuring dynamic shard rebalancing and automated failure recovery from network partitions.
- **Logging-Based FS**: Implemented a persistent, crash-recoverable file system with in-memory and on-disk logging, granular sync/abort, and optimized I/O via cached reads, differential writes, and semaphore-based concurrency control.
- **Fair-Share Scheduler & QoS Compression**: Built a Xen-like fair share credit-based thread scheduler with load balancing achieving 2-3× performance improvement and a QoS-enabled multi-threaded compression daemon using message queues, shared memory, and worker thread pools for concurrent file processing with client fairness guarantees.
- **Scalable LLM-Guided Adaptive Fuzzing Framework**: Engineered a hybrid testing pipeline utilizing LLM-based fuzz-harness generation and vulnerability prediction for adaptive resource allocation, achieving 85.7% precision and a 1.7x speedup in detecting algorithmic bugs over state-of-the-art baselines.
- **Edge AR Streaming Analysis**: Designed a WebRTC AR streaming system in Pion and FFmpeg with multicast pipelines and per-frame hooks to measure FFmpeg filter versus Mediapipe overlay overheads, then benchmarked on Jetson TX1.
- **Analogical Reasoning Evaluation for LLMs**: Benchmarked LLM analogical mapping on story-based tasks using embedding similarity and two-step prompting, uncovering gaps in causal reasoning and sensitivity to paraphrase variation.
- **C-Inspired Language Compiler**: Designed a C-like language with jagged arrays, operator overloading, and type-safe operations. Built a recursive-descent parser to generate ASTs and perform semantic analysis with efficient memory usage.

Technical Skills

Languages and Frameworks: C/C++, GoLang, Python, Java, Bash, PyTorch, React, SQL
Developer Tools: Git, Linux, Neo4j, AWS, gRPC, PostgreSQL, MongoDB, Ansible, Docker, Kubernetes, etcd, patroni
Teaching Assistant: Advanced Operating Systems (Graduate level), OOP (Undergraduate level)