Chandra Kiran Narala

linkedin.com/in/chandrakirannarala | chandrakirannarala@gmail.com | 1-551-331-7492 | New York, NY

SUMMARY

Performance-focused Software Engineer with a Master's in Electrical Engineering from NYU Tandon. Specialized in building real-time data processing systems and hardware-accelerated solutions. Proven track record in developing high-throughput pipelines and ultra-low latency architectures. Passionate about applying data-driven engineering to high-performance racing environments.

EDUCATION

NYU Tandon | MS in Electrical Engineering | USA — Jan 2022 - May 2024

Courses: Computer System Architecture, Advanced VLSI, Advanced Hardware Design, DSP

GVPCOE | Bachelors in Electronics and Communications Engineering | India — Jul 2016 - Sept 2020

Courses: Data Structures & Algorithms, System Design & Optimization, Computer Architecture, Digital Signal Processing, OS

PROFESSIONAL EXPERIENCE

SRAM DV Engineer | NYU, USA — Jan 2024 - Nov 2025

- Led a 4-person research team to optimize a 256×4-bit SRAM array in 7 nm FinFET, using mathematical modeling and statistical analysis to reduce power consumption by 20% compared to 6T SRAM cell
- Developed and verified SRAM testbenches using Cocotb, enabling Python-based verification with functional and latency testing
- Executed Timing Analysis to achieve timing closure while optimizing performance, area, and power trade-offs. Resolved critical timing violations for reliable read and write operations at higher clock frequencies, optimizing memory access stability
- Collaborated with physical design team on RTL-level fixes during timing closure, hands-on exposure to Cadence ASIC flow

Software Engineer | Cognizant Technology Solutions, India — Nov 2020 - Dec 2021

- Designed and developed high-throughput Python-based systems handling 100K+ requests/sec for real-time data processing, improving system throughput by 15% and reducing latency by 20%
- Built distributed pipelines enabling low-latency data ingestion and transformation for operational decision-making
- Implemented monitoring, logging, and debugging strategies to maintain 99.9% uptime in production environments
- Developed fault-tolerant systems using retry logic and circuit breakers under high-load conditions
- Collaborated with QA, software, and business teams across the full SDLC, from requirements clarification and implementation to testing, deployment, and production support

PROJECTS

Python Real-Time Transit Data Aggregation Service

- Designed and deployed a distributed backend service integrating real-time MTA data across 7 feeds and 22 subway lines; Built high-throughput data processing pipeline with sub-second latency using Python
- Implemented API routing and request optimization reducing redundant calls by 85%; Deployed production-ready service with fault tolerance, monitoring, and zero-downtime restarts

C++-Based Huffman Compression System

- Implemented C++ system using data structures and algorithms for lossless compression with memory management and efficient I/O
- Developed a high-performance file compression tool implementing Huffman coding algorithm for optimal data storage solutions
- Achieved 47% size reduction with 1.89:1 compression ratio and sub-second processing for 2MB files. Delivered perfect lossless reconstruction with 0.28s encoding and 0.37s decoding performance

Biometric Fingerprint Identification System

- Developed a biometric fingerprint identification system using Python and OpenCV that processes raw fingerprint images through a 3-stage pipeline including background removal, ridge enhancement, and SIFT feature extraction
- Implemented FLANN-based matching algorithm with geometric verification using RANSAC, achieving greater than 90% accuracy with 15+ keypoint correspondences and reducing false positives by 40%
- Built complete end-to-end solution with Tkinter GUI for real-time processing, processing comparisons in ~4 seconds per fingerprint

Distributed Ultra-Low-Latency Message Parsing Engine

- Engineered a performance optimized message parsing engine (C++/ Verilog) on FPGA for ultra-low latency FIX parsing in HFTs
- Designed a custom FSM-driven design featuring a round-robin FIFO buffer for parallel message storage and precompiled BRAM lookup tables for rapid tag=value mapping and real-time validation, thereby eliminating the runtime overhead of XML parsing
- Projected to achieve up to 200 times faster processing; delivering latencies in 50µs range and handling millions of messages per second compared to CPU-based parsing, with 7ms average latency and maximizing throughput in real-time data environments

SKILLS

Programming: C++ (Modern C++), Python (PyTorch, NumPy, Pandas, SciPy), C#, HTML, CSS, JavaScript
Development Tools: Git workflows, Linux, SQL, Cocotb/Pytest for verification,
Hardware/Performance: Verilog, SystemVerilog, VHDL, FPGA systems, Xilinx Vivado, Matlab, Cadence Simulation tools
Systems & Methodologies: OOPs, Microservices Architecture, Simulation systems, Distributed Systems, Agile / Scrum, Data analysis
Protocols & Techniques: Ethernet (TCP/IP), system-level optimization, System Design, Monitoring and logging, SDLC