srik.palani@gmail.com
(612) 513-3787

Srikar Palani

linkedin.com/in/srikar-palani/
Maple Grove, MN

Education

University of Illinois Urbana-Champaign — Champaign, IL
M.S. in Computer Science — *May 2027*
B.S. in Computer Engineering, ***GPA: 3.81/4.0***

Coursework: Operating Systems, Networking, Parallel Programming, Applied Machine Learning

Honors: James Scholar, Dean's List

Skills

Languages: C++, Python, C, Java, SystemVerilog, RISC-V Assembly, JavaScript, HTML/CSS, Dart

Technologies: AWS (Lambda, API Gateway, Neptune, CDK, CloudFormation, S3, VPC, IAM), CUDA, React, Next.js, FastAPI, Flutter, PostgreSQL, Supabase, Git, Pandas, PyTorch, OpenCV, Vivado, Tailwind CSS

Work Experience

Meta — Menlo Park, CA
Software Engineering Intern — RecSys - Inference - Serving — *May 2026*

- Architected and implemented a deterministic execution framework that transforms declarative JSON specifications into reproducible SEV hypothesis validation pipelines, replacing non-deterministic AI-driven incident investigations
- Designed a domain-specific JSON specification and modular execution engine with typed interfaces, schema validation, batching, retries, versioning, content hashing, and persistent caching to automate telemetry analysis over Scuba and ODS
- Reduced incident investigation time from up to two weeks to approximately one day (>10× faster), expanded automated validation from individual models to thousands of production models (∼1000× coverage), and supported analysis applicable to 75%+ of historical SEVs

Amazon Web Services — Seattle, WA
Software Development Intern — Infrastructure Security — *May 2025 – August 2025*

- Built and deployed a production platform managing millions of network devices, centralizing network visibility and access control across Amazon's global infrastructure for network engineers
- Designed and implemented Python backend APIs using AWS Lambda, API Gateway, and Neptune to automate ACL validation, access path verification, and role-based access control
- Developed a React interface for interactive visualization of production network topologies, deployed through CloudFront to provide low-latency global access

Infocrunch — Remote
Software Development Intern — *Jun 2024 – Aug 2024*

- Developed Flutter and React applications supporting workforce management across 3,500+ rural banks, including attendance, transaction, account, and payout tracking
- Built a FastAPI backend using Pandas, PostgreSQL, and Supabase to process Excel uploads, normalize data, and manage cloud authentication and storage

Projects

UNIX-like RISC-V Operating System:

- Built a Unix-like operating system on RISC-V, implementing virtual memory, demand paging, system calls, and user/kernel isolation
- Developed a custom filesystem with block-level caching and VIRTIO-backed storage for efficient file operations
- Implemented a preemptive scheduler and inter-process communication via pipes, enabling concurrent process execution

GPU-Accelerated Convolution (CUDA, Tensor Cores):

- Implemented optimized convolution kernels using CUDA WMMA (TF32) on NVIDIA A40 GPUs
- Fused compute operations to reduce memory overhead and improve convolution throughput
- Optimized shared memory usage, warp tiling, and parallel execution for improved GPU performance