NATHAN SAMSON

nsamson4@umbc.edu | Baltimore, MD, USA | linkedin.com/in/nathan-samson-bostesa | nathan-tes-samson.com/

EDUCATION

University of Maryland - Baltimore County **August 2023 – May 2027**
Bachelor's, Computer Science *GPA: 3.57*

PROFESSIONAL EXPERIENCE

Capital One **Mclean, VA, USA**
Software Engineer Intern (TEIP) *June 2025 – August 2025*
- Built a serverless synthetic monitoring solution leveraging Python, AWS Lambda, and Fargate to execute automated health checks across production services, supporting both cron-based and API-triggered validation workflows.
- Architected containerized FastAPI applications with Docker, implemented comprehensive APM using New Relic, and streamlined deployment processes through AWS CLI automation and Serverless Framework configurations.
- Extended the platform with PagerDuty alerts, S3 HTML reports, and a results UI to speed incident triage and historical analysis.

University of Maryland - Baltimore County **Baltimore, MD, USA**
Undergrad researcher *September 2023 - Present*
- Co-authored PSMark, a distributed pub/sub benchmarking framework implementing 12 IoT workloads from 7 real-world datasets (smart cities, factories, healthcare) in Erlang to evaluate MQTT/DDS systems at scale. Submitted to IEEE Peercom 2026. **Link to repo**
- Co-authored and architected MQTT-DAP, a privacy-preserving MQTT extension enabling GDPR compliance through purpose-based access control. Modified Eclipse Mosquitto (60K LOC) with ~3K lines of C to enforce data subject rights (erasure, rectification, access) at the protocol level. Submitted to ACM SenSys 2026. **Link for broker. Link for benchmark**
- Designed and benchmarked 5 purpose-management approaches (broker-modifying and broker-agnostic) and 3 operational categories while maintaining sub-second latency for sensitive data operations.
- Built a Smart Campus occupancy prediction system integrating LiDAR sensors with ML models to forecast dining hall congestion. Engineered edge-cloud data pipelines with MQTT, optimized SQL/TSDB storage for high-velocity streams
- Developing an LLM-powered IoT simulation pipeline that generates synthetic meeting scenarios and multi-modal sensor data (temperature, CO_2, energy) to validate smart-building models.

OmniSyncAI **Remote**
Software Engineer Intern *May 2024 - July 2024*
- Engineered user-friendly CRM account setup using Node.js, React, and PostgreSQL, reducing onboarding time and increasing team invitations through AI-powered recommendations
- Built backend API endpoints to handle user authentication, team management, and account provisioning workflows
- Implemented data models and database schemas in PostgreSQL to support multi-tenant CRM functionality

PROJECTS & OUTSIDE EXPERIENCE

Embedded lightweight vector database for serverless & edge - *Link to project*
- Built lightweight vector database in Go with Python bindings optimized for serverless achieving sub-millisecond performance (0.05ms inserts, 0.1ms lookups, 2ms similarity search) with 2MB footprint
- Designed memory-mapped storage engine with hash indexing enabling embedding caching to reduce OpenAI API costs through get-or-compute patterns

Formally Verified Code Generation - *Link to project*
- Built LLM code generator using Claude Sonnet 4 API with Dafny formal verification, achieving 92% first-try success on LeetCode Easy-Medium problems with mathematically proven correctness
- Architected multi-stage verification loop (Parse Spec → Generate Python → Generate Dafny → Verify → Refine) handling recursion, binary trees, and dynamic programming with theorem proving

Multi-Modal Product Search System - *Link to project*
- Engineered unified commerce agent handling text, image-based, and multimodal search using CLIP embeddings (ViT-B/32), FAISS vector search, and MMR diversity guarantees
- Built deterministic intent router with strict retrieval policies, contextual elimination, hallucination reduction through rules-based filtering, and metadata filtering (brand, category, price)
- Designed comprehensive evaluation framework

SKILLS

Languages: Python, Go, C/C++, JavaScript, Java, Erlang, SQL
ML/AI: TensorFlow, PyTorch CLIP, FAISS, NLP, Prompt Engineering, Claude/Gemini APIs
Infrastructure: Docker, Kubernetes, AWS (Lambda, Fargate), FastAPI, Node.js, React/Next.js, PostgreSQL
IoT/Systems: MQTT, DDS, Edge Computing, Distributed Systems, New Relic

ACHIEVEMENTS

- Nvidia Summer Bridge Program Participant
- Capital One Tech Summit Participant
- UMBC CSEE Research Day 2024 Best poster Award