Mathew Bino

+ | https://www.linkedin.com/in/mathew-bino/
mathewmyid@gmail.com | https://github.com/machuii

Professional Summary

Software Dev Engineer with over a year of experience building and operating scalable data systems processing billions of records daily. Passionate about solving complex technical problems while delivering measurable customer impact through innovative solutions.

Skills

Languages:	Python, C, C++, SQL, Bash
Operating System:	Ubuntu 22.04 LTS
Cloud & Infrastructure:	GCP (Cloud Run, GKE, PubSub, BigQuery), AWS (EC2), Azure (Service Bus, Blob Storage), Docker, Kubernetes, Snowflake,PySpark
Frameworks, Libraries & Tools:	FastAPI, Django, Pillow, PyTorch, Pandas, Matplotlib, OpenCV, NiFi, Git, Lex, Yacc, Gcc, Gdb,Spark

Education

- B.Tech Computer Science and Engineering, National Institute of Technology Calicut (NIT-C), Kerala | CGPA:8.85 | Graduated: 2025
- 12th Board CBSE (98.2 %) | JEE MAINS: 3546 rank |Kerala State Entrance: 107 rank

Work Experience

Micron Technology ,Data Engineer, SMAI (July 2025 - present)

Backend and Data Engineer with Yield Management Team, responsible for end-to-end orchestration of data from global fabrication, assembly and test sites. Managing the ingestion and transformation in all forms of data ranging from csv, parquet and json to live camera feeds. Maintain pipelines that integrate factory tool telemetry and test results to drive real-time manufacturing yield insights.

MapViewer Backfill

(PySpark, BigQuery, Snowflake , GCS)

Rebuilding a year of semiconductor test history (die/wafer/package diagnostics) in BigQuery.

- Cut a PySpark/Dataproc ETL job from 161 to 75 minutes wall-clock .Identified a config change that eliminated memory spill zero and doubled peak concurrency
- Delivered a 1.5Tb global deduplication in a single 6 hour Spark job .
- Observed Snowflake result retrieval making a significant contribution to total time so migrated the Snowflake table to BQ. Landing 13 billion rows in 2 hours

API Migration & Infrastructure Change - SMAPI GKE

(FastAPI,Python, GKE, Docker,HashiCorp Vault,Cloud Trace)

- Owned and executed FastAPI version upgrade across multiple repositories, resolving breaking changes and dependency resolution to ensure service stability.Leveraged Cloud Trace to diagnose high response latency for credential retrieval, implementing caching that improved performance.
- Transitioned to the GKE Gateway API switching to ClusterIP from LoadBalancer centralizing SSL certificate handling and better management of multiple instances.

- Strengthened security through robust HashiCorp Vault integration and transitioned from LDAP to AppRole-based authentication.Switched from hardcoded ADC to using rotated GCP credentials, identifying and resolving an API framework compatibility bug in the process.

HBM Tool Data Ingestion
(Python, Snowflake, PubSub, Scheduler, Workflows, EventArc, Cloud Run)
Tool data of both engineering and production recipes, measuring in billions of records per day come in from factory sites.

- Implemented Snowflake performance optimizations , including the use of internal staging, asynchronous execution, and COPY INTO commands for efficient bulk loading.Conventional write_pandas approach used was taking 75 minutes for 124 Million rows and was reduced to 8 mins with these optimizations.

TTE(Test to End) Pipeline Migration Support
(GKE , PubSub , BigQuery , Bigtable,Bash)

TTE is Micron's global chip-testing data collection system, processing over a million files daily to help engineers determine chip quality for shipping

- Engineered a reprocessing pipeline that reduced daily missing summaries from 3,000+ to under 100. Integrated a reusable Cloud Function to standardize batch routing to downstream flows (MFZ, TTE-Extractor, AVRO), centralizing logic and eliminating redundant processing.
- Authored a real-time TTE monitoring dashboard (7 views: live pipeline, backload status, missing-batch investigation, Pub/Sub backlog health,5-stage lot-trace tool and a resubmission page) giving engineering and operations end-to-end visibility and self-service reprocessing
- Developed an automated hourly monitoring script for parquet generation;successfully detected a GCP switchover failure, ensuring system reliability and proactive stakeholder communication

InFLY (Intelligent Forecasting and Learning of Yields)
(Github,Cloud Run Job , Workflows)

Micron data-science and engineering platform that produces and continuously self-corrects semiconductor yield forecasts used for production planning and financial goal-setting. It combines an ML "formulator" feedback-loop engine with a data pipeline, a set of API services, and a UI.

- Successfully orchestrated the migration of a C#-based backend to a fully operational FastAPI architecture within a three-week timeframe.
- Developed a bespoke workflow that empowers users to initiate curve reprocessing based on altered inputs through the user interface.
- Built a stateful GitHub CD setup leveraging file fingerprinting and delta matching to streamline initial production provisioning and automate incremental code deployments for 42 components.

Federated Learning (Ongoing)
Successfully implemented two swarm based Google's federated learning algorithms, FedPSO and FedCPSO on an open-source framework, Flower as part of B.tech Project. Additionally, optimized the existing FedCPSO algorithm, achieving a 70% reduction in on-device memory usage and a 49% decrease in network communication costs . These improvements enhance the feasibility and efficiency of federated learning training on resource-constrained edge devices.
Paper presented at **ComNet 2026 (IEEE Xplore)**. Forthcoming publication.

Paid Intern, Resume Ranking and Classification, Calpine Technologies (**May - Jul 2024)**

NLP, SentenceTransformers, SentenceSimilarity, NER

Algorithm for ranking resumes to match a given job description. Used the techniques of natural language preprocessing, Sentence-Transformers and Sentence similarity metrics. The tool has gone live now. Works on any resumes at the rate of 3 Resumes per sec. Created a backend API to use the model. Additionally worked on a multi class classification model utilizing Bert output that can help categorize resumes. The model achieved an accuracy of 98.74%. - partially automated the process of acquiring data by scraping LinkedIn using Selenium.

Paid Intern, PixDynamics, Infopark, Kochi (May – Jun 2023)

Python, OpenCV, Google Colab

An AI based project with the objective to detect tables from an individual's resume. Client was a naval shipping company. Object extraction and structure detection using YOLOv5, fine-tuned on a custom dataset .

Information extraction using LayoutLMv2, based on the word piece tokenizer. I also worked with OpenCV, Pillow, LabelImg for dataset preparation

Curriculum/Hobby projects

eXpOS (Jul - Dec 2023) – Designed and implemented a full-featured operating system from scratch including process synchronization, semaphores, user management, file management, paging, interrupts and multi-core support. Demonstrated deep understanding of OS fundamentals and low-level system programming.

Simple Programming Language(Jan - May 2024) - Designed and built a programming language from scratch using Lex, Yacc and C. with inheritance, subtype polymorphism, type definitions and function implementation. Applied compiler design principles and formal language theory to create functional language specification

Self Hosted OpenVPN - A simple openvpn setup on AWS EC2 coupled with an elastic IP

Certifications

- Machine Learning with Apache Spark - IBM
- Agentic AI with LangChain and LangGraph - IBM