Michael Medford, Ph.D

I am an astrophysicist that loves investigating, organizing and scaling big data to solve intractable problems. I have applied these skills to a diverse set of challenges ranging from scaling R&D products on cloud infrastructure to constructing calibration pipelines for the world's largest telescopes.

MichaelMedford@gmail.com · 973.600.0340 · www.MichaelMedford.com · github.com/MichaelMedford

Skills

SQL, Airflow, Snowflake, BigQuery, ELT Pipelines, Kafka, Data Modeling & Warehousing, Database Design, Python, Go, C, Ruby on Rails, CircleCI, Codefresh, Git, Docker, Kubernetes, Terraform, AWS, GCP
Data Architecture, Data Visualization, Big Data Analytics, Agile Project Management, Technical Writing

Work Experience

Staff Software Engineer - Subject Matter Expert in Datastores
Insight M · July 2024 - Present

- Developed a two-year organizational roadmap to improve data architecture and security, including a staging and development data strategy, a data transformation pipeline for enhanced business intelligence dashboards, and a data retention overhaul, saving the company six figures annually
- Optimized database processes by refactoring high-cost computational components, reducing runtime expenses, and enhancing system resiliency with comprehensive test coverage
- Enhanced data infrastructure by implementing a database event alert pipeline, launching an internal application for query plan analysis to improve query efficiency, and introducing the company's first postmortem template to foster learning from software retrospectives

Staff Software Engineer - Systems Base
Aumni, a JPMorgan Chase Company · July 2023 - June 2024

- Redesigned our backend architecture with management, product and engineering to enable flow in stream-aligned teams
- Developed a migration plan to maintain 100% uptime while refactoring our Ruby on Rails monolith into a set of independent services within the JPMorgan ecosystem
- Constructed contract testing framework to ensure independent deployability in collaboration with our Testing Enablement Team

Senior Software Engineer - Data Chapter
Aumni, a JPMorgan Chase Company · November 2022 - June 2023

- Led and executed refactor of company's data processing model to unlock cross-functional analytics
- Designed, documented and deployed to production an independent microservice for indexing millions of documents into OpenSearch in only 6 weeks
- Invented an SQL unit test framework on top of a custom file parsing library that runs in Snowflake and executes via Airflow pipelines validating our core business logic
- Overhauled the infrastructure and tooling for developer data synchronization to reduce sync time from 15 minutes to 20 seconds

Technical Lead of Planet Fusion Monitoring Pod
Planet Labs · March 2022 - October 2022

- Architected 10x increase in our daily Python processing while cutting per unit costs 80% in four months
- 98% reduction in database query latency & CPU utilization via SQL schema & query optimization
- Cut GCS & AWS data delivery processing overhead by 30x using batched GCP pubsub messages
- Moved team into a bi-weekly sprint structure with ticket triage, sprint retrospectives and sprint planning

Geospatial Software Engineer
Planet Labs · June 2021 - October 2022

- Implemented automated data backups & disaster recovery via terraform to comply with ISO certification
- Set up on-call notifications by integrating Sentry, PagerDuty and Slack into our Python & Go stack
- Created a comprehensive cost estimate method to measure profit margin goals for our product teams

Part-time Geospatial Software Engineer
Planet Labs · November 2020 - May 2021

- Enabled end-to-end development by connecting our React front-end, Go orchestrator & Python scripts
- Conducted our largest scale test to date, giving the green light to onboard new customer contracts
- Transitioned change detection algorithm to parallelized cloud context for a 100x speedup in E2E runtime

Research
Experience

A New Method for Detecting Solar System Objects on High Performance Computers
Lawrence Berkeley National Laboratory · January 2016 - September 2019   w/ Dr. Peter Nugent

- Invented planet detection pipeline that searched 100+TB of images using 20,000+ lines of Python and C
- Implemented real-time neural network scoring of planet candidates in Dockerized Python Flask apps
- Engineered HPC scheduler to execute 1,000+ compute processes via many-to-many SQL databases
- Measured accuracy and completeness by applying statistical methods to artificially injected signals

Detecting Black Holes in the Milky Way using Simulations and Observational Analysis
University of California: Berkeley · August 2018 - June 2020   w/ Prof. Jessica Lu

- Predicted event rates for hypothetical telescope surveys by executing galaxy simulations
- Designed OOP solution to include new astrophysical phenomenon in Bayesian model fitting process
- Reduced pipeline execution time by 50% through memory profiling and IO optimization

Parallelized Executable for Removing Noise from Telescope Images
Lawrence Berkeley National Laboratory · September 2019 - June 2020   w/ Dr. Peter Nugent

- Constructed physical models of atmospheric fringes in optical images with principle component analysis
- Built parallelized feature identification and extraction tool currently running on 50,000+ images per night
- Released code as an open source Python package: *fringez*   fringez Documentation

Education

PhD, Astrophysics, University of California: Berkeley   May 2021
    Advisors: Jessica Lu and Peter Nugent
    Thesis: *Discovery of Rare Signals in Large Scale Time Domain Surveys: Dark Planets and Black Holes*

MA, Astrophysics, University of California: Berkeley   GPA: 3.79 / 4.00 | 2017

BS, Physics and Astronomy, Northwestern University   GPA: 3.75 / 4.00 | 2011
    Weinberg College of Arts and Sciences
    Advisor: Dr. Michael Smutko, Collaborator: Dr. Vicky Kalogera
BS, Theatre, Northwestern University   2011
    School of Communication

Public
Software
Packages

PopSyCLE (2020)   PopSyCLE Documentation
Population Synthesis for Compact-object Lensing Events

- Developed pipeline infrastructure to execute code in high performance supercomputing environments
- Reconfigured data format schema to extendible compound HDF5 to enabling additional image filters
- Introduced code reviews and unit tests to collaboration workflow among five person team

zort (2019)   zort Documentation
ZTF (Zwicky Transient Facility) Object Reader Tool   Sole Author

- Executed spatial cross-matching and filtering for time-domain measurements of billions of objects
- Official data reader for ZTF telescope Public Data Releases, representing 11+ international institutions

Awards

UC-National Lab In-Resident Graduate Fellowship
The Regents of the University of California   04/2019 - 03/2021

- Awarded $130,000 to discover isolated black holes at the Lawrence Livermore National Laboratory

Interests   Astrophotography, Racquetball, Chess, Theatre, Financial Coaching