Deerfield, IL 60015

Andrew Repp

andrew@andrewrepp.com

andrewrepp.com

(708) 476-3829

Employment

Software Engineer at VMware — June 2022 - Present

- Developed key features within backup and restore utilities for Greenplum database. Notably includes parallel metadata restore, improving restore of schema DDL by up to 2,000% in more complex customer systems. See article. (Go)
- Architected and implemented from-scratch rewrite of backup manager utilities, reducing runtime latency by approximately 8,000% on very large customer deployments. (Go, sqlite)
- Designed and developed new SQL command (ALTER TABLE REPACK) to improve compression and scan performance on very large customer datasets. Observed 50% reduction in disk space usage in tests of large installs. (C)
- Implemented several key features in core Greenplum server codebase, to help drive to release of GPDB7. Examples include autovacuum implementation and parallelized system monitoring views. See video tutorial. (C)
- Dedicated GitHub profile for my open source contributions at VMware

Software Engineer at Network Perception — April 2021 - June 2022

- Profiled and improved performance of core backend, inference, and parser code. Resulted in approximately 10,000% improvement in runtime performance. (Python, Redis)
- Re-architected and implemented system parallelization, leveraging kernel APIs and python wrapper APIs, to dramatically improve both backend runtime performance and frontend responsiveness. (Python, Linux)
- Designed and implemented a full rebuild of the product's integrated analytics engine. (Python, JavaScript)
- Developed parser and inference integration for various networking software output formats, including Cisco ASA firewalls, Fortilink Fortiguard firewalls, Google Cloud Platform, AWS Networking, Nessus, nmap, and libpcap. (Python)

Senior Data Scientist & Team Lead at OptionCare Health — July 2019 - April 2021

- Developed software package for analysis of patient medication adherence. Authored internal whitepaper showing company excellence relative to industry standards, submitted to conference. (Python, SQL)
- Developed schema and ETL queries for analytic data mart, leveraging existing primary data sources into an enriched platform for rapid analysis, centralizing the business on aligned methodologies and definitions for all metrics. (SQL)

Senior Analyst & Manager at OptumRx — January 2017 - June 2019

- Developed statistical programs deployed to visualization layer hosted on department website. Owned every step of multiple projects using this workflow, from requirements to deployment. (SAS, Tableau, Azure)
- Scoped, architected, developed, and deployed a new algorithmic approach to optimally categorize products by price to minimize cost to the company and maximize patient satisfaction (SAS, SQL).
- Refactored company reporting suite into a highly performant parallel architecture, reducing runtime by 400%.

Notable Technical Projects

- OSS Medication Adherence Package (2020). Developed and published package for calculating medication adherence. Used pandas, NumPy, and multiprocessing. Contracted as a consultant on its use and validation for multiple private healthcare entities. Ported to rust with clap and rayon. Discussed here. (Python, Rust)
- Formulary Optimization System (2019). Developed and deployed a program to analyze medication usage across 65 million patients using customized algorithm and grid search. Productized by employer. (SAS, SQL)

Education

Georgia Institute of Technology	Master of Science - Computer Science
University of Wisconsin (Stevens Point)	Master of Science - Fisheries Science
University of Illinois (Champaign-Urbana)	Bachelor of Science - Natural Resources

Languages and Technologies

- Go (2 years); Python (5 years); SQL (10 years); SAS (7 years, Base SAS Certified); Bash (4 years); C/C++ (3 years).
- Git (8 years); Linux (6 years); Redis (2 years); Agile Process (4 years).