

Albert Miller

New York, New York | 405-420-5832 | albert.miller.nyc@gmail.com

Education

The University of Texas at Austin | August 2016 - December 2020 | *Bachelor of Science in Mechanical Engineering with a Business Administration minor*

Skills

Languages: Python, C++, Bash Scripting, SQL, Scala

Cloud Services: Certified AWS Solutions Architect

Tech: Apache Cassandra, Apache Spark, Valgrind, Pandas, Boto, Linux, Snowflake, Docker, Jenkins, Splunk, Git

GitHub: <https://github.com/albertcmiller1>

Work Experience

Capital One | Cloud Data Engineer | August 2022 - Present | *Big Data Pipeline Team*

- End of day batch job processing for Capital One's financial core
- Wrote spark software in scala to extract auto loan transaction data from cassandra database, transform data into both producer and consumer schemas, and upload to data lakes via aws step function, EMR, lambda, and S3 buckets
- Developed on a validation repo using python pandas to validate data from upstream to downstream
- Performance testing and tuning of spark configurations to speed up our extraction jobs
- Built aws lambda function in python to upload S3 parquet files to data lakes
- Large data generation using python pandas (thousands of gigs) for performance testing of our pipeline with consideration on data model so partitions will evenly distribute across the database and data frames
- Built dashboards to view and monitor pipeline hardware and software logs with splunk, new relic and aws sns

Capital One | Software Engineer | August 2021 - August 2022 | *SRE Cassandra Cloud Database Team*

- Wrote database infrastructure recovery software using python and bash scripts to recover an apache cassandra database of n servers from their aws snapshots. Currently used by Capital One's biggest databases
- Worked closely with new relic to debug why their process was saturating our server's thread counts and overloading our garbage collectors, eventually crashing our production cassandra servers
- Integrated open source software into our cassandra environment with cloud deployment automation

Projects

Financial Exchange | C++

- Order book hosted on an aws EC2 server hosting a websocket and api allowing users to place limit orders
- Three threads continuously posting buy and sell orders to stimulate the book and create liquidity
- Project gave the opportunity to practice data structures, recursion, memory management, multithreading, and sockets

Trading Platform | Python

- Command line tool to act as a broker to the order book (building out GUI next)
- Place trades, visualize the live spread, view your portfolio and transactions, and plot or stream stock data history
- Use of pandas, plotly, aws api gateway, lambda, and dynamo database to record user's transactions

Personal

I have found a love for C++ and Python and am seeking an opportunity to practice and improve on these skills. Outside of coding, I am an 8 handicap golfer and I received my private pilot's license in college. I also love fishing, ping pong, and billiards.