Khang Ly

khangly@berkeley.edu
linkedin.com/in/khangly
(626) 618-6969

Education

- **University of California - Berkeley** — Berkeley, CA
  *Bachelor of Science, Electrical Engineering and Computer Sciences; GPA: **3.98*** — *Jun. 2018 – Dec 2020*
  - **Coursework**: Data Structures, Operating Systems, Algorithms, Database Systems, Computer Architecture, Parallel Programming, Computer Security, Artificial Intelligence, Machine Learning, Signals and Systems, Probability, Optimization Models, FPGA Digital Design

Skills

- **Languages (in order of proficiency)**: Python, C, C++, Java, Go, SQL, Scheme, PHP, JavaScript, Shell Script
- **Technologies**: AWS, Docker, Kubernetes, Flutter, Spark
- **Tools**: Bazel

Experience

- **Theorem LP** — San Francisco, CA
  Software Engineer Intern — *Jun. 2020 – Aug. 2020 (8 weeks)*
  - Theorem buys loans from lending marketplaces using a model they has developed.
  - Made an extract, transform, load (ETL) pipeline that scores all the loan data (about 5M) in the data warehouse.
  - Redesigned the ETL pipeline to use a fixed amount of memory so they won't be killed by the Out of Memory Killer. Reduced memory usage of some ETLs up to 90%.
  - Sped up the process of loading external data to Theorem's data warehouse from days to minutes by making data schema inference and parquet conversion from running locally to running on the cloud using AWS Glue.
  - **Technologies**: Bazel build tool, Docker, Kubernetes, Amazon Redshift, AWS Glue (PySpark)

Personal Projects

- **Hot Deals Alert** — github.com/khangly/HotSD
  - Built a web service to identify hot deals.
  - Deals data are collected using a Python script from the website slickdeals.net. Every minute, the script is triggered by `cron` to download and parse the contents using BeautifulSoup, then append the new data to a MySQL database.
  - All data points are aligned based on the time each deal is posted and during the first hour. Each feature is graphed as a time series, and is fitted by an exponential curve. A deal is "hot" if all of its features lie above the curves.

Class Projects

- **Database Management System (DBMS)** — Spring 2020
  - Implemented B+ Tree and various join algorithms. Added an optimizer that searches and executes a low estimated cost query based on some heuristics.
  - Finally, implemented multigranularity locking so queries can run concurrently.
- **End-to-End Encrypted File Sharing System** — Fall 2019
  - Designed and implemented a secured file sharing system in Go that guarantees the confidentiality and integrity of data stored in an **untrusted** server.
  - The project was ranked second among 200 groups.
- **Pint Operating System (PintOS)** — Summer 2019
  - Implemented multi-level feedback queue scheduler.
  - Implemented various file operation syscalls, such as `create`, `remove`, `open`, `read`, `write`.
  - Finally, added a buffer cache to improve performance of read and write of recent files and expanded PintOS file system to handle extensible files and subdirectories.

Awards

- **Jim and Donna Gray Endowment Award** — UC Berkeley
  Recipient — *Spring 2019 - Fall 2019*
  - Awarded to computer science students completing their junior year who have demonstrated both high scholastic achievement and financial need.