林子軒 Lin Sam · ✉ samlin266118@gmail.com · ☎ 0972724528 · in sam lin · ◯ linsamtw ·

Summary

7-8 years of experience with data engineer and soft engineer. (Distributed Queue System, RMDB, NoSQL, TSDB, Web Crawling, RESTful API, ETL, Docker Swam, GKE(K8S), CICD, GCP, Airflow ...etc.)
The Github project gets 2,000 stars.

Work Experience

17 LIVE - Data Engineer Manager   May. 2021 - Now

- **Reduce data team 50% cost** by building a BQ monitoring dashboard and optimizing Pub/Sub, DataFlow, Composer, GCE, GKE, BQ storage, and BQ slots.

- The main developer for refactoring ETL. Creating an airflow project by Cloud Composer to transfer ETL tools from digdag to airflow and transfer ETL develop method from shell script to python. (More than 20 large pipelines.)

- Maintenance BigQuery more than 300 tables and more than 10 TB data.

- Create pipelines from the backend team, multiple mysql and mongodb, and appsflyer to bigquery.

- Create Data Team's first real-time ETL system via GKE, Pub/Sub, and Memorystore for sending push notifications to users. (Graceful Shutdown is also set.)

- Create a good development culture, including the introduction of CICD, dev-stage-uat-master, release news, unit tests, and test coverage.

- Using Airflow unified scheduler job, like cloud function scheduler, bq scheduler, crontab, and ML model by R or Python...etc.

- Create Data Team's first real-time ETL system via GKE, Pub/Sub, and Memorystore for sending push notifications to users.

- Create Data Team's first API via GKE for ML model, include achieve graceful shutdown, run stress test via ApacheBench, and setup auto-scaling by hpa. 95% latency is less than 200 ms, and RPS is more than 200.

- Keep SLO above 99%.

SinoPac Holdings - Software Engineer   Nov. 2019 - May. 2021

- Develop **python, C#** api for stock/option/future trading.

- Deploy a test system to simulate trading.

- Collecting distributed system log by **elk**, **Grafana** and **Prometheus**, 13 GB log data/day.

Tripresso - Data Engineer   Oct. 2018 - Oct. 2019

- Analysis travel data and build a machine learning model. Estimating increase **3%** orders (revenue).

- Maintain and develop an **distributed queuing system for web crawling** with 20 machines.

Publications

- Published a book related to data engineering. Python 大數據專案 X 工程 X 產品資料工程師的升級攻略.

Projects

FinMind python project   **2,000 Stars On Github.** Sep. 2017 - Now

- Create a payment flow using **ECPay**, and get some order from the user.

- More than 8,000 are registered.

- Using **vue, python and django** to develop web pages. (https://finmindtrade.com/).

- The API has more than **600,000** requests per day. Upper bound of api is **8,000/minute** request by ApacheBench.

Education

National Dong Hwa University (GPA : 3.87)	**Tamkang University**
Master of Science, Sep. 2017.	Bachelor of Science, Sep. 2015.
Major : Mathematics and Statistics.	Major : Mathematics.