

Hao-Yu Shih

: <https://www.linkedin.com/in/hao-yu-shih/>

: +1-608-692-3281 : haoyushih25@gmail.com : <https://github.com/hshih25>

Education

University of Wisconsin-Madison
M.S. in Computer Sciences

Sep. 2021 - May 2023
Madison, WI

National Cheng Kung University
B.S. in System and Naval Mechatronic Engineering

Sep. 2015 - Jun. 2019
Tainan, Taiwan

Skills

Programming Languages	Java, C++, C, Python, Erlang, Rust, PHP, JavaScript, Haskell
Software/Frameworks	PyTorch, Spark, Hadoop, gRPC, Thrift, OpenCV, MySQL, Node.js, Nginx
Build/DevOps Tools	Git, Mercurial, make, Docker, AWS

Experience

Meta May 2022 - Aug. 2022
Software Engineer Intern (Erlang, C++, Python, Thrift, MySQL | Backend) *Menlo Park, CA*

- Built new loggings for the Data team could fetch the data efficiently to facilitate data analysis and evaluate a new pricing model.
- Built a new logging to enhance visibility of Webhook information and reduced the time needed to resolve tickets by **80%**.
- Implemented an auditing feature using Python, MySQL, data pipeline, and internal tools to alarm the team when the guarantee which WhatsApp give to the business customer is not achieved.
- Compared the data between WhatsApp core database and WhatsApp Business Manager database with Python and MySQL to prevent the inconsistency between databases.

Projects

Leftover-Killer: an Android based recipe application (PHP, Nginx, MySQL | Backend)

- Built the backend system with PHP, Nginx, and MySQL.
- Implemented backend API executor code in PHP for server to execute.
- Implemented the Regression test and Unit test by PHPUnit.
- Increased the backend code coverage to **100%**.
- Designed the ER Diagram to construct the Database.

Big Data System (Python, PyTorch, Spark, Hadoop | Machine Learning)

- Deployed and configured distributed ML Learning frameworks on PyTorch in conjunction with Gloo, a MPI like communication framework.
- Implemented the PageRank algorithm, like what Google uses to evaluate the quality of links to a webpage, on Apache Hadoop as the underlying file system and Apache Spark as the execution engine.

Replicated Block Storage (C++, gRPC, Primary/Backup | Distributed System)

- Built a strong-consistent and high-available network-based block storage system with Primary/Backup strategy.
- Implemented the system with C++ and used gRPC to create the channels between each servers.
- Designed server states: Primary, Backup and Recovery states to prevent the system crash and help the servers recover from the crash.
- Designed data states: Committed, Dirty, and Pending-Commit to prevent the corruption and data loss.

SDN and Router Implementation (Java, TCP, SDN | Computer Network)

- Implemented a layer-3 routing application which will install rules in SDN switches to forward traffic to hosts.
- Developed a distributed load balancer to redirect new TCP connections to hosts in the round-robin order.
- Implemented the ICMP, ARP, and RIP behaviours in the virtual router.
- Built a routing table by using Distance Vector Routing.