Cai, Yifan

Email: fyc1007261@live.com | Homepage: http://www.yifancai.tech

EDUCATION

Shanghai Jiao Tong University   Shanghai, China
Bachelor of Engineering, Software Engineering, System Software   Expected in June 2020

> **Overall GPA:** 3.77 / 4.30 (88.19/100); **Major GPA:** 3.88 / 4.30 (89.95/100)
> **Honors/Awards:**
> - SJTU Scholarships (2016-2017, 2017-2018 & 2018-2019)
> - Huawei Scholarships (2017-2018)
>
> **Relevant Courses:** Operating Systems, Distributed Systems, Computer Architecture

PUBLICATIONS

Measuring the Effect of Data Center Resource Disaggregation on Production DBMSs [Experiments and Analyses]. (Under Review)

> Qizhen Zhang, **Yifan Cai**, Xinyi Chen, Sebastian Angel, Ang Chen, Vincent Liu, and Boon Thau Loo.
> Submitted to *46th International Conference on Very Large Data Bases (VLDB), 2020*

Rethinking Data Management Systems for Disaggregated Data Centers.

> Qizhen Zhang, **Yifan Cai**, Sebastian Angel, Ang Chen, Vincent Liu, and Boon Thau Loo.
> To appear in *Conference on Innovative Data Systems Research (CIDR)*, 2020

Consensus-based Data Statistics in Distributed Network Systems. [Paper]

> **Yifan Cai**, Jianping He, Wenbin Yu, and Xinping Guan.
> In the *57th IEEE Conference on Decision and Control, Miami, USA,* December 2018

RESEARCH EXPERIENCE

University of Pennsylvania   PA, US
Research Intern, Advisor: Dr. Vincent Liu   Jul 2019 – Dec 2019

- Conducted research on the impact of disaggregated data centers in the design of relational databases
- Developed microbenchmarks of nested loop join, hash join, and grace hash join operators and investigated the relationship between local memory size, the number of remote memory access and the performance degradation of each operation
- Added features such as relative paths to LegoOS (a disaggregated operating system) by adding more system calls in order to deploy complex systems such as PostgreSQL and MonetDB on it.
- Analyzed the performance drawbacks of existing disaggregated operating systems for query executions, and outlined potential solutions
- Categorized database queries by the operators used and analyzed to which degree resource disaggregation would lead to the performance degradation of each operator

Lab of System Control and Information Processing, SJTU   Shanghai, China
Research Assistant, Advisor: Dr. Jianping He   Sep 2017 – Jun 2019

- Invented a protocol in distributed network systems to compute probability density functions in a fully distributed way and enable multiple compute nodes to share their statistics
- Designed and optimized two probability distribution functions (PDF) to calculate algorithms which run under both public and anonymous network conditions

University of Michigan   MI, US
Research Intern, Advisor: Dr. Xianglei Huang   Aug 2018 – Sep 2018

- Built an LSTM-based solar forecasting system using the Keras deep learning library
- Designed the prediction model with an approximate error of only 10%, which was less than all other prediction models in published papers based on the same dataset

PROJECTS

Distributed File System   Shanghai, China
A Project of Computer System Engineering Course, SJTU   Sep 2018 – Jan 2019
- Developed a distributed file system with Fuse that could support basic Linux interfaces (touch, move, write, unlink, etc.) and work with multiple clients
- Implemented distributed locks with the local system cache to achieve atomicity and high performance
- Implemented replication protocols using heartbeats and notifications in case of random failures to improve the availability of the whole system

Tiger Language Compiler   Shanghai, China
A Project of the Compilers Course, SJTU   Sep 2018 – Jan 2019
- Developed a compiler in C for the *Tiger* language to generate assembly code, following the x86-64 calling conventions
- Optimized the compiler by conducting liveness analysis and register allocations

Smart Electric Appliance Monitoring and Scheduling System   Shanghai, China
A Project of the Project Management and Software Development Course, SJTU   Jun 2018 – Jul 2018
- Collaborated with team members to build a smart electric appliance monitoring and scheduling system, which could collect information from all sensors and operate the appliances
- Developed the backend of the web application in Java and integrated the hardware with C++ and Python,
- Integrated a series of cross-process communications for smart scheduling of appliances, deep learning based solar energy forecasting, and convenient human-computer interactions, and ensured the consistency among them
- Exhibited in the SJTU Software Exhibition, 2018

B+ Tree Based Database   Shanghai, China
A Project of the Data Structure Course, SJTU   Jul 2017 – Aug 2017
- Developed a B+ tree-based database system which could support CRUD (create, read, update and delete) operations at a speed of 1M+ operations per second when running with over 1M rows of data
- Designed and implemented a buffer system between the disk and memory to improve the performance of reading and writing data

SKILLS

Programming Languages: C/C++, Python, Java, JavaScript, SQL
Languages: Mandarin (Native), English (TOFEL 111)
Certificates: Deep Leaning Specializations (by Coursera), Machine Learning (by Coursera), The Bits and Bytes of Computer Networking (by Coursera).