Zeyuan Hu

Homepage: https://zhu45.org/
Email: ferrishu3886@gmail.com

EDUCATION

University of Texas **Austin, TX** **Sept 2017 – May 2019**
- M.S. in Computer Science. (GPA: 3.88/4.00)

University of Wisconsin **Madison, WI** **Sept 2010 – Dec 2014**
- B.A. in Computer Science. (GPA: 3.74/4.00)
- B.A. in Economics with Honors. (GPA: 3.85/4.00)
- B.A. in Mathematics. (GPA: 3.81/4.00)

WORK EXPERIENCE

Cloud Architect Engineer **State Street Financial Service** **June 2019 – Current**
Omnia storage team
- Developing auto-deployment system of IBM Cloud Object Storage in multi-site clusters using Ansible, Docker
- Leading development of a distributed workload generator and performance benchmark toolkit written in Go

Software Engineer Internship **Schlumberger** **May 2018 – August 2018**
HPC infrastructure team
- Implemented a monitoring component of the in-house High-Performance Computing (HPC) engine in C++ to provide the fault tolerance and handle the "straggler" problem
- Employed SGD algorithm to dynamically learn the best timing for backup executions of the in-progress tasks based on the computation task characteristics
- Built a C++ code generator that automatically generates the application layer code based on the engine API

Software Engineer **IBM** **August 2015 – August 2017**
Db2 LUW federation team
- Constructed Hive and Impala wrappers with C++ and Java to support federation database between traditional RDBMS and Hadoop-based data warehouse solution
- Created automated setup tools with Shell that reduce product configuration time by 75%
- Enhanced server option optimization tools using C to reduce federation database performance tuning time by 90 % and enable the capability of tuning the product against Hive, Impala, and Spark
- Resolved over 20 defects, including a severe memory leak issue that impacted a $1.6 million deal

SPECIALIZED SKILLS

- **Languages:** C++, C, Python, Go, Rust, Shell, SQL, Java, Elisp, MATLAB
- **Software:** CMake, Autotools, QEMU, Docker, Tensorflow, Keras, Git, ClearCase, Hive, Impala, Hadoop
- **Graduate Coursework:** Machine Learning, Structured Models for NLP, Human Computation & Crowdsourcing, Natural Language Processing, Semantics, Distributed Systems, Advanced Operating Systems, Data Centers, Algorithms, Automated Logic Reasoning,

PUBLICATION

- Jialin Wu, **Zeyuan Hu**, Raymond J. Mooney. "Joint Image Captioning and Question Answering" In *VQA Challenge and Visual Dialog Workshop at the 31st IEEE Conference on Computer Vision and Pattern Recognition* (CVPR 2018)
- Jialin Wu, **Zeyuan Hu**, Raymond J. Mooney. "Jointly Generating Captions to Aid Visual Question Answering". (ACL 2019 *Oral*)

SELECTED PROJECTS

- **RustFS** (2018 -). Building a user-space file system that leverages NVMe SSD. `Rust`, `SPDK`
- **Strata with Lease** (2018). Extended Strata file system with Lease mechanism to support concurrent file access across processes. `C`
- **HyperPebblesDB** (2018). Constructed a key-value store that is part of LevelDB family with focus on reducing write amplification. `C++`, `CMake`, `Autotools`
- **Distributed Key-Value Store** (2018). Built a distributed key-value store with `Python` that uses *eventually consistency* model with two session guarantees: *Read Your Writes* and *Monotonic Reads*.

HONORS AND AWARDS

- 2018 **Best Internship Project Award (Software Engineering)**, Schlumberger
- 2017 **IBM Appreciation program for the Practice: Dare to create original ideas**, IBM
- 2016 **IBM Manager's Choice Award - Put the Client First**, IBM
- 2016 **IBM China Development Laboratory Hackathon - 2nd Place**, IBM
- 2014 **Graduation with Distinction**, University of Wisconsin
- 2013 **Honors Summer Sophomore Research Apprenticeship**, University of Wisconsin
- 2012 **Meek Bishop Scholarship in Economics**, University of Wisconsin
- 2010-2012 **Dean's List**, University of Wisconsin

TEACHING

- CS386D Database Systems (Spring 2019, UT-Austin). Teaching Assistant
- EE382V Data Engineering (Fall 2018, UT-Austin). Teaching Assistant
- NEU466M Quantitative Methods in Neuroscience (Spring 2018, UT-Austin). Teaching Assistant
- M408K Differential Calculus (Fall 2017, UT-Austin). Teaching Assistant

MANUSCRIPT

- **Zeyuan Hu** and Julia Strout. Exploring Stereotypes and Biased Data with the Crowd. arXiv preprint arXiv:1801.03261 (2018)

RESEARCH EXPERIENCE

Research Assistant **UT-Austin** **August 2018 – Present**
Prof. Vijay Chidambaram
- Building a user space file system on top of NVMe SSD leveraging `SPDK` library from Intel and `Rust`

Research Assistant **UT-Austin** **April 2018 – December 2018**
Prof. Emmett Witchel
- Used a cloud service benchmark (*CloudSuite*) to measure the performance penalty brought by encryption on the IPC between a web server (Nginix) and a PHP application in a local environment
- Measured journaling impact on the write amplification of various file systems using *filebench* (customized), *blktrace*, *iostat*, and *strace*
- Measured the write amplification of workloads from *filebench* and Git workload from *BetrFS* on file systems that are fragmented (i.e., *age*), which is created from file system aging tool (e.g., *Geriatrix*)

Research Assistant **UW-Madison** **May 2013 – April 2014**
Prof. Vikas Singh
- Applied Spatial Gaussian Process & Dirichlet Process on fMRI data with `MATLAB` and improved power of testing on predicting Dementia based upon pixel value of the scan by 5 %

SERVICE AND SOCIETIES

- UTCS Master Admission Committee (Jan 2018 – March 2018), Member
- IBM Diamond & Ring Toastmaster Club (Jun 2016 - Jun 2017), President