Willis Wang

Research Interests

Distributed Systems, Databases, Computer Systems

Education

2018–Present **B.A., Computer Science**, *University of California, Berkeley*.
GPA: 3.72/4
Coursework: Graduate Computer Systems, Operating Systems, Database Systems, Compilers, Networking, Computer Architecture, Computer Security, Algorithms, Data Structures

Research Projects

2021–Present **Dataspread**, *Berkeley RISE Lab*, Prof. Aditya Parameswaran.

- Dataspread is a spreadsheet system with databases as the back-end, enabling an interactive and intuitive view alongside high performance and scalability for data.
- Built on internal infrastructure to fix bugs in cell clearing and provide core features for spreadsheet management.
- Currently developing an out-of-order scheduler for cell updates, maximizing the usability of cells in user view while maintaining overall performance.
- **Technologies and Libraries**: Java, PostgreSQL

2021–Present **PSL Multicast**, *Fog Robotics*, Prof. John Kubiatowicz.

- Paranoid stateful lambdas (PSL) provide a function-as-a-service utility that allows execution of trusted code with confidential data on potentially untrusted edge devices with eventually-consistent semantics.
- Designed and implemented a multicast message-passing protocol for paranoid stateful lambdas on the Global Data Plane, reducing network congestion while maintaining security predicates.
- Contributed to writing core router and message passing code as well as testing and benchmark code.
- Connected existing Paranoid-SGX codebase to ns-3 for network simulation.
- Held a poster session and wrote a research-style paper on findings.
- **Technologies and Libraries**: C++, Protobuf, ZeroMQ, ns-3

2020 **Delta Lake Caching**, *Databricks*.

- Delta Lake, although offering high performance in large and complex workloads, suffers from slow point-lookup and write operations.
- Worked with Databricks engineer Burak Yavuz on a write-back caching layer using RocksDB for Delta Lake, enabling OLTP workloads on top of Delta Lake.
- Personally designed and wrote core RocksDB cache layer, benchmarking tools, and test workloads.
- Achieved over 10x speedup in single-row accesses for cache hits.
- **Technologies and Libraries**: Scala

2020 **COVID-19 Contact Tracing**, *VIP Lab*, Prof. Avideh Zakhor.

- A production-level COVID-19 contact tracing application, using smartphones' magnetometer sensors to detect proximity in order to preserve user privacy.
- Personally contributed to writing the simulation framework, writing scripts to parse and transform online datasets, and training a binary classifier for detecting extended contact.
- **Technologies and Libraries**: Python, Pandas, AutoML, Azure ML

Industry Positions

Summer 2021 **Software Engineer Intern**, *Amazon*, Seattle, WA.

- Lead design and implementation of favorites system for internal AWS account management.
- Deployed application to 150k+ end users within Amazon's Consumer Division.
- **Technologies and Libraries**: Javascript, AWS Lambda, API Gateway, DynamoDB

Summer 2020 **Software Engineer Intern**, *8th Wall*, Palo Alto, CA.

- Architected storage, backup, retention, and restoration system for 10k+ customer repositories.
- Automated database migrations, rollback, and CDN updates across all internal engineering infrastructure.
- **Technologies and Libraries**: Python, Javascript, AWS Lambda, S3, API Gateway, DynamoDB, Cloudwatch, Cloudfront, SNS, Elastic Beanstalk

Teaching

Spring 2022 **CS162 Undergraduate Student Instructor**, *UC Berkeley EECS*, Berkeley, CA.

- *Operating Systems and Systems Programming*, Profs. John Kubiatowicz and Anthony D. Joseph

2021 **CS162 Course Staff**, *UC Berkeley EECS*, Berkeley, CA.

Teach students in conceptual and technical topics through office hours and review sessions, develop on PintOS project infrastructure, manage grading for student work.

- **Fall 2021 CS162 Reader**, *Operating Systems and Systems Programming*, Prof. Ion Stoica
- **Spring 2021 CS162 Reader**, *Operating Systems and Systems Programming*, Profs. Anthony D. Joseph, Natacha Crooks

Spring 2020 **CS182 Course Staff**, *UC Berkeley EECS*, Berkeley, CA.

Contributed to rubric and grading for homeworks, tests, and projects for the class.

- **CS182 Reader**, *Designing, Visualizing and Understanding Deep Neural Networks*, Prof. John Canny