# Qinzhe Wu

**Electical and Computer Engineering**
**University of Texas, Austin**
**qw2699@utexas.edu**
**(512)806-8824**
**https://bambowu.github.io**

## Research Interests

Parallel processing architecture design, especially interested in optimizing workloads with task-level parallelism via software-hardware co-design approaches.

## Education

**Ph.D. Program**, Architecture Computer System Embedded System 05/2023
**Dissertation**: Architectural Support for Message Queue Task Parallelism
University of Texas, Austin; GPA 4.00 Austin, TX
**Bachelor of Engineering**, Information Engineering 06/2017
Zhejiang University (ZJU); GPA 3.89 Hangzhou, China

## Courses

| | | | |
|---|---|---|---|
| **Distributed Systems** | Distributed computation, Consensus | A | Fall 2020 |
| **Multicore Computing** | Parallel algorithms, Synchronization | A | Spring 2020 |
| **Datacenters** | Networking, Microservice, RDMA, RPC | A | Fall 2019 |
| **Superscalar Microprocessor** | VLIW, Multi-issue, Branch prediction | A | Spring 2019 |
| **Security at the HW/SW Interface** | Side Channel Attack, Memory Integrity | A | Fall 2018 |
| **User System Interplay** | System Design, Research Paper Writing | A | Fall 2018 |
| **Performance Evaluation** | Workload Characterization, Profiling | A | Spring 2018 |
| **Advanced Embedded MCU System** | Kernel, Driver, HW/SW Co-design | A | Spring 2018 |
| **Computer Architecture** | LC3b, Simulator, Caches, VM, OoO | A | Fall 2017 |
| **Advanced Programming Tools** | Python, Web, Andriod | A | Fall 2017 |
| **Computer Organization and Design** | MIPS, Pipeline, Memory Hierarchy | 94/100 | Fall 2015 |
| **Digital System Design** | Logic Circuit, FPGA | 99/100 | Spring 2015 |

## Projects

**Speculative Push for Anticipating Message Requests in Multi-Core System** 01/2022
Implemented a message queue speculation scheme to reduce the cross-core communication latency
**Proxy Benchmark Development for Data Warehouse Workloads** 12/2021
Developing a open-source benchmark to represent data query engine workloads in datacenters
The speculation algorithm achieve 1.33× speedup over the state-of-the-art hardware queue
**Algorithm for Byzantine Lattice Agreement in Synchronous Message Passing Systems** 11/2020
Designed an algorithm to solve Byzantine Lattice Agreement by $3log(f) + 3$ synchronous rounds
**Architectual Support for Multi-Producer-Multi-Consumer Communication** 10/2020
Designed an architecture to assist messages passed from producer core to consumer core faster
Evaluation on 6 benchmarks shows 2.09× speedup on average and 61% less memory traffic

**Machine Learning Workload Characterization** 06/2019

Similarity and dissimilarity analysis on 3 ML benchmark suites (MLPerf, DAWNDench, DeepBench)

Profiling-guided scheduling for multi-GPU multi-model training achieves 1.16× speedup

**Graph Data Placement Optimization on a Heterogeneous Memory System** 04/2019

Extended Ligra graph processing framework for heterogeneity-aware data placement

Testing various of decision strategies based on heuristics from graph topology

**A Study on Anti-Spectre Binary Analysis** 12/2018

Studied and evaluated oo7, a plugin to Binary Analysis Platform for Spectre path search

Came up with new test cases as well as modification to recognize more variants (BTI and SSB)

**Accelerator for Graph Processing** 05/2018

Designed accelerators on Zynq®-7000 for PageRank

DMA supported high bandwidth BRAM for high throughput with deep pipeline

12.7x speedup over naive software implementation for 100 iterations on graphs with 384 vertices

**Distributed Graph Processing System with Redundancy Reduction** 02/2018

Experiments with 4 frameworks on 10 graphs were conducted for comparisons

16.3× speedup was achieved on average by avoiding unneccessary computations

**Performance Proxies For SPEC CPU 2017** 11/2017

Used Perf to sample workload performance for phase level behaviors analysis

Conducted instrumentation with Intel® PIN tool

Synthesised proxies based on the workload profiles to mimic performance

**MIMO Processing Based on Zynq®-7000**, Degree Project, Advised by Dr. Jianyi Yang 05/2017

Built up a testbed for multi-mode ONoC MIMO processing

**Web Server Based on STM32 and ESP8266** 11/2016

Developed a daemon under $\mu$C/OS to response HTTP requests

Dynamically provided web browser with files in SD card managed by FatFs (SSI and CGI used)

**Radio-controlled Car Based on STM32 and RF24L01**, Group Leader 07/2016

Three-person group designed a car model moving along tracks drawn on touch screen

Designed a fancy UI and implemented a circular FIFO to buffer handwring data

**Multi-Cycle Processor Using BASYS3** 01/2016

Up to 69 instructions in the Intel 8051 ISA were supported by the design

Ran a LED digital clock program (compiled by Keil from C to binary code) on the processor

## Work Experience

Configurable Benchmark for Data Serving Workload, *Internship* Summer 2021

**Facebook Inc.** Menlo Park, CA

Based on fbthrift, built a server/client benchmark framework, which can be configured to plug in various of computation components in order to represent data serving workloads.

Automated Workload Similarity Analysis, *Internship* Summer 2020

**Facebook Inc.** Menlo Park, CA

Developed an automated framework to evaluate how similar two workloads are. The framework reports similarity scores and highlights the difference on 7 pairs of proxy benchmarks and target applications, and helps to improve one of the benchmarks.

Virtual Link, *Internship* Summer 2019

**Arm Research** Austin, TX

Designed Virtual Link, a light-weight synchronization mechanism, and wrote a cycle-accurate simulator to evaluate the core part.

Exascale Computing Project, *Advanced Short Term Research Opportunity* Summer 2018

**Oak Ridge National Laboratory Future Technologies Group** Oak Ridge, TN

Characterized exascale workloads running on GPU, and extended ASPEN (Abstract Scalable Performance Engineering Notation) models for GPU applications.

Multi-channel Cam Controller Development, *Part-time* 06/2017
**Hangzhou Huicheng Automation System Co., Ltd.** Hangzhou, China
Developed a product that detects the spin then switchs cams on/off accordingly. The controller supports up to 12 channels and communicates with HMI panel via Modbus protocol for configuration.

## Teaching Experience

**EE382N Performance Evaluation and Benchmarking**, *Teaching Assistant* Spring 2021
University of Texas at Austin Austin, TX USA
I assisted Prof. Lizy John in grading, and mentored a few course projects, such as MLPerf Kernel Extraction, Deep Learning Workloads Characterization, and Quantify Worload Similarity.

**Introduction to Computing System**, *Teaching Assistant* 07/2017
Zhejiang Univerisity, Summer Course Hangzhou, China
Topics include the logic circuit, LC-3 (Little Computer 3) architecture, LC-3 assembly language, data types, system service routine, memory-mapped I/O, interruption, function calls. I assisted Prof. Yale Patt in grading, recitations, and lab mentoring.

## Publications

**SPAMeR: Speculative Push for Anticipated Message Requests in Multi-Core Systems**, Qinzhe Wu, Ashen Ekanayake, Ruihao Li, Jonathan Beard, Lizy K. John (ICPP22)

**Virtual-Link: A Scalable Multi-Producer, Multi-Consumer Message Queue Architecture for Cross-Core Communication**, Qinzhe Wu, Jonathan Beard, Ashen Ekanayake, Andreas Gerstlauer, Lizy K. John (IPDPS21)

**Hot Regions in SPEC CPU2017**, Qinzhe Wu, Steven Flolid, Shuang Song, Junyong Deng, Lizy K. John (IISWC18)

**Hardware-aware 3D Model Workload Selection and Characterization for Graphics and ML Applications**, Ruihao Li, Aman Arora, Sikan Li, Qinzhe Wu, Lizy K. John (ISQED22)

**Demystifying Graph Analytics Frameworks and Benchmarks**, Junyong Deng, Qinzhe Wu, Xiaoyan Wu, Shuang Song, Lizy K. John (SCIS20)

**Accelerating Force-directed Graph Layout with Processing-in-Memory Architecture**, Ruihao Li, Shuang Song, Qinzhe Wu, Lizy K. John (HiPC20)

**Demystifying the MLPerf Training Benchmark Suite**, Snehil Verma, Qinzhe Wu, Bagus Hanindhito, Gunjan Jha, Eugene B. John, Ramesh Radhakrishnan, Lizy K. John (ISPASS20)

**A Study of Core Utilization and Residency in Heterogeneous Smart Phone Architectures**, Joseph Whitehouse, Qinzhe Wu, Shuang Song, Eugene John, Andreas Gerstlauer, Lizy K. John (ICPE19)

**Metrics for Machine Learning Workload Benchmarking**, Snehil Verma, Qinzhe Wu, Bagus Hanindhito, Gunjan Jha, Eugene B. John, Ramesh Radhakrishnan, Lizy K. John (FastPath19)

**Start Late or Finish Early: A Distributed Graph Processing System with Redundancy Reduction**, Shuang Song, Xu Liu, Qinzhe Wu, Andreas Gerstlauer, Tao Li, Lizy K. John (VLDB19)

**Experiments with SPEC CPU 2017: Similarity, Balance, Phase Behavior and SimPoints**, Shuang Song, Qinzhe Wu, Steven Flolid, Joseph Dean, Reena Panda, Junyong Deng, Lizy K. John

## Awards

➤ Outstanding Gradutes of Zhejiang University, ZJU (Class of 2017 Undergraduates) 2017
➤ First-Class Scholarship for Outstanding Students, ZJU (TOP 3%) 2016

| | |
|---|---|
| ➤ “Zhebao-Ali” Scholarship, Zhejiang Daily Press & Alibaba (4/182) | 2016 |
| ➤ 3rd Prize of “TP-LINK Cup” College Students Electronic Design Contest | 2016 |
| ➤ Honorable Winner, Mathematical Contest in Modeling, COMAP | 2016 |
| ➤ Outstanding Student Leader Awards, ZJU | 2015 |

## Skills

**Laguage:** C/C++, Python, Java, Verilog, Perl, Assembly, LaTeX, MATLAB, Javascript, HTML, CSS
**Software:** VIM, Git, Gem5, Intel® PIN and Vtune, Perf, Snipersim, NVIDIA DIGITS, numba, nvprof, Google Cloud Platform, Vivado, Modelsim, Keil, Altium Designer, Multisim, Andriod Studio