

# Xuhao Luo

9500 Gilman Drive, La Jolla, CA 92093

x3luo@eng.ucsd.edu

## Education

---

**University of Illinois Urbana-Champaign**

Aug 2021 - Now

Ph.D. Student in Computer Science

Advisor: Prof. Tianyin Xu

**University of California San Diego**

Sep 2019 - Mar 2021

M.S. in Computer Science

Advisor: Prof. Yiyang Zhang

GPA: 3.82/4.0

**University of Science and Technology of China (USTC)**

Sep 2015 - Jun 2019

B.S. in Applied Physics

Major in Microelectronics and Solid State Electronics

## Research Interests

---

Operating System, Networking, Computer Architecture, Heterogeneous Computing

## Research Experience

---

**An Asynchronous Executor for Distributed ML System**

Jun 2020 - Sep 2020

*Research Project at Microsoft Research Asia*

- Designed and implemented an asynchronous executor for task scheduling and dispatching on multiple hardware.
- Designed and implemented CUDA-based high-performance inter-GPU communication channel for distributed ML within large-scale GPU cluster.
- Multi-GPU collective operation(AllReduce, AllGather, Broadcast) throughput outperforms Nvidia NCCL by at most 18.4% under the same system setting.

**An FPGA-based Disaggregated Memory System**

Sep 2019 - Dec 2020

*Research Project at UCSD (In Submission)*

*Supervisor: Prof. Yiyang Zhang*

- Working on FPGA-based disaggregated virtual memory system for system resource disaggregation.
- Designed and implemented a go-back-N based full reliable network stack on both FPGA and host Linux server to support high-performance reliable network communication. Using kernel-bypass to achieve high-throughput and low-latency.
- Designed and implemented an RPC-semantic connectionless network stack to improve scalability, with a delay-based congestion control.
- Achieved RDMA-like latency and 10Gbps(limited by hardware interface) throughput at rack-scale.

## Research Publication

---

Zhiyuan Guo\*, Yizhou Shan\*(co-first author), Xuhao Luo, Yutong Huang, Yiyang Zhang, **Clio: A Hardware-Software Co-Designed Disaggregated Memory System (Preprint)**

## Past Projects

---

**Design and Implementation of HLS Based Quantized Neural Network Accelerator**

Jan 2019 - May 2019

*Graduation Project*

*Supervisor: Prof. Xi Jin*

- Studied the 8-bit quantization algorithm, including the quantization algorithm, the dequantization algorithm and the implementation of the 8-bit quantized convolution.

- Designed and implemented a general 8-bit quantized convolution module on Xilinx Virtex FPGA, which achieved high parallelization through array architecture, and realized memory access optimization through data reuse.
- Developed the TensorFlow C++ API for the hardware accelerator using OpenCL. Used this accelerator to accelerate the ResNet-50 CNN and achieved a speedup of 5.17x and a memory usage reduction of 66% compared with the CPU TensorFlow implementation on Xeon E5 2686.

### Binary Neuron Network (BNN) Acceleration using HLS

Jul 2018 - Sep 2018

*Summer Internship at Cornell University*

*Supervisor: Prof. Zhiru Zhang*

- Designed and implemented a BNN accelerator for LeNet-5 for MNIST handwritten digits recognition.
- Applied multiple methods to improve the performance of the accelerator including parallelization, pipelining, line buffer, task-level parallelism and batch processing.
- Implemented the accelerator on Zedboard, ZC706, and AWS EC2 F1. Achieved speedups of 33x(580fps), 88x(1543fps) and 114x(2170fps) compared with the software implementation baseline on Intel Xeon 5420 CPU.

### Honors and Awards

- |   |          |
|---|----------|
| • USTC Class of 2019 Outstanding Graduates  | May 2019 |
| • 2017/18 USTC Outstanding Students Scholarship, Golden Award                             | Sep 2018 |
| • 2016/17 USTC Outstanding Students Scholarship, Silver Award                             | Sep 2017 |
| • 2015/16 USTC Outstanding Students Scholarship, Bronze Award                             | Sep 2016 |
| • The 13 <sup>th</sup> Competition of Physical Research Experiment, 2 <sup>nd</sup> Prize | Dec 2017 |
| • The 6 <sup>th</sup> Aegon-Industrial Fund Scholarship                                   | Jun 2017 |

### Skills

---

<b>Language</b>	C/C++, Python, Go, Rust, Haskell, OpenCL, Verilog, HTML, JavaScript
<b>Tools/Framework</b>	TensorFlow, Docker, Zookeeper, LLVM, Google Test