

JIACHENG MA

Ph.D. Candidate

Computer Science and Engineering
University of Michigan

2260 Hayward Street
Ann Arbor, MI 48109, USA

jcma@umich.edu
<https://jcma.me>

RESEARCH INTERESTS

My research is at the intersection of hardware and software. I am interested in improving the programmability, reliability, debuggability, and deployability of heterogeneous systems by building systems support such as hypervisors, compilers, debuggers, and runtimes. I am also interested in system virtualization and software-hardware co-design.

EDUCATION

University of Michigan

Ann Arbor, MI, USA

Ph.D. Candidate in Computer Science and Engineering

Sept. 2018 – Present

Thesis Topic: Systems and Debugging Supports for Reconfigurable Hardware

Advisor: Prof. Baris Kasikci

University of Michigan

Ann Arbor, MI, USA

M.S.E. in Computer Science and Engineering

Sept. 2018 – 2021

Shanghai Jiao Tong University

Shanghai, China

B.E. in Software Engineering

Sept. 2014 – June 2018

Thesis: Efficient GPU Live Migration Optimized by Software Dirty Page for Full Virtualization

Advisor: Prof. Zhengwei Qi

PEER-REVIEWED CONFERENCE PUBLICATIONS

- [1] Debugging in the Brave New World of Reconfigurable Hardware. Jiacheng Ma, Gefei Zuo, Kevin Loughlin, Haoyang Zhang, Andrew Quinn, and Baris Kasikci. *Proceedings of the 27th International Conference on Architectural Support for Programming Languages and Operating Systems*, 2022.
- [2] MEGATRON: Software-Managed Device TLB for Shared-Memory FPGA Virtualization. Yanqiang Liu, Jiacheng Ma, Zhengjun Zhang, Linsheng Li, Zhengwei Qi, and Haibing Guan. *The 58th Design Automation Conference*, 2021.
- [3] DOLMA: Securing Speculation with the Principle of Transient Non-Observability. Kevin Loughlin, Ian Neal, Jiacheng Ma, Elisa Tsai, Ofir Weisse, Satish Narayanasamy, and Baris Kasikci. *30th USENIX Security Symposium (USENIX Security 21)*, 2021.
- [4] Execution reconstruction: Harnessing failure reoccurrences for failure reproduction. Gefei Zuo, Jiacheng Ma, Andrew Quinn, Pramod Bhatotia, Pedro Fonseca, and Baris Kasikci. *Proceedings of the 42nd ACM SIGPLAN International Conference on Programming Language Design and Implementation*, pages 1155–1170, 2021.
- [5] A Hypervisor for Shared-Memory FPGA Platforms. Jiacheng Ma, Gefei Zuo, Kevin Loughlin, Xiaohe Cheng, Yanqiang Liu, Abel Mulugeta Eneyew, Zhengwei Qi, and Baris Kasikci. *Proceedings of the 25th International Conference on Architectural Support for Programming Languages and Operating Systems*, 2020.

- [6] gRemote: API-Forwarding Powered Cloud Rendering. Dongjie Tang, Yun Wang, Linsheng Li, Jiacheng Ma, Xue Liu, Zhengwei Qi, and Haibing Guan. *Proceedings of the 29th International Symposium on High-Performance Parallel and Distributed Computing*, pages 197–201, 2020.
- [7] gMig: Efficient GPU Live Migration Optimized by Software Dirty Page for Full Virtualization. Jiacheng Ma, Xiao Zheng, Yaozu Dong, Wentai Li, Zhengwei Qi, Bingsheng He, and Haibing Guan. *Proceedings of the 14th ACM SIGPLAN/SIGOPS International Conference on Virtual Execution Environments*, pages 31–44, 2018.
- [8] gScale: Scaling up GPU Virtualization with Dynamic Sharing of Graphics Memory Space. Mochi Xue, Kun Tian, Yaozu Dong, Jiacheng Ma, Jiajun Wang, Zhengwei Qi, Bingsheng He, and Haibing Guan. *Proceedings of the 2016 USENIX Conference on Usenix Annual Technical Conference*, pages 579–590, 2016.

PEER-REVIEWED JOURNAL PUBLICATIONS

- [9] gRemote: Cloud rendering on GPU resource pool based on API-forwarding. Dongjie Tang, Linsheng Li, Jiacheng Ma, Xue Liu, Zhengwei Qi, and Haibing Guan. *Journal of Systems Architecture*, 116:102055, 2021.
- [10] gMig: Efficient vGPU Live Migration with Overlapped Software-based Dirty Page Verification. Qiumin Lu, Xiao Zheng, Jiacheng Ma, Yaozu Dong, Zhengwei Qi, Jianguo Yao, Bingsheng He, and Haibing Guan. *IEEE Transactions on Parallel and Distributed Systems*, 2019.
- [11] Scalable GPU Virtualization with Dynamic Sharing of Graphics Memory Space. Mochi Xue, Jiacheng Ma, Wentai Li, Kun Tian, Yaozu Dong, Jinyu Wu, Zhengwei Qi, Bingsheng He, and Haibing Guan. *IEEE Transactions on Parallel and Distributed Systems*, 29(8):1823–1836, 2018.

PEER-REVIEWED WORKSHOP PUBLICATIONS

- [12] Efficient Runtime Mercurial Core Detection with Core-Specific Test Case Synthesis. Jiacheng Ma, Gefei Zuo, Andrew Quinn, and Baris Kasikci. *1st Workshop on Data Integrity and Secure Cloud Computing (DISCC'22)*.
- [13] Tolerate Silent Data Errors with Coded Computation. Gefei Zuo, Jiacheng Ma, Andrew Quinn, and Baris Kasikci. *1st Workshop on Data Integrity and Secure Cloud Computing (DISCC'22)*.

U.S. PATENT

- [14] Jiacheng Ma, Haibing Guan, Zhengwei Qi, and Yongbiao Chen. System, apparatus, and method for optimizing a scalable GPU virtualization, October 1 2019. US Patent 10,430,991.

RESEARCH EXPERIENCE

Efes Lab, advised by Prof. Baris Kasikci

University of Michigan

Hardware Defects Detection

Sept. 2022 – Present

- This on-going work explores techniques to construct software test cases that detects silicon defects. Once finished, it would enable data center operators to detect hardware failures in a cost-efficient way.

FPGA Record/Replay

Nov. 2021 – Present

- This on-going work explores record/replay techniques on FPGAs. In this work, we build the framework that can not only record and replay a hardware execution on an FPGA, but also enable other developers to build more sophisticated development tools.

- FPGA Bug Study and Debugging Techniques* [1] Sept. 2020 – Nov. 2021
- This work explores debugging techniques for FPGAs. In this work, we perform a comprehensive study on real-world bugs in FPGA projects, classify these bugs based on their root causes and symptoms, and propose techniques to help bug localization.
- Secure Speculative Execution* [3] Jan. 2020 – Apr. 2020
- This work presents the first defense scheme to provide automatic comprehensive protection against all known transient execution attacks.
- Hardware-Assisted Bug Reproduction* [4] Aug. 2019 – Apr. 2020
- This work combines online recording and offline symbolic execution to recover failing program executions with low overhead and high accuracy. In the online phase, the control-flow of a program execution is recorded; the control-flow is then used in the offline phase to guide symbolic execution.
- FPGA Virtualization* [5] Apr. 2018 – Aug. 2019
- In this project, we create the first scalable hypervisor for shared-memory FPGA platforms. The hypervisor, called OPTIMUS, supports both spatial multiplexing and temporal multiplexing, and scales linearly until the memory bandwidth is saturated. OPTIMUS helps data center operators to deploy multiple shared-memory accelerators on the same FPGA, thus maximizing the resource utilization.
- TCloud Lab**, advised by Prof. Zhengwei Qi *Shanghai Jiao Tong University*
- Remote Rendering* [6, 9] Mar. 2018 – May 2018
- This work enables API-forwarding based cloud rendering for OpenGL applications on a resource pool.
- Optane as Memory* (with Intel) Aug. 2017 – June 2018
- This work explores using NVM (Intel Optane) as main memory for KVM guests.
- vGPU Live Migration* [10, 7] (with Intel) Dec. 2016 – July 2017
- Intel GVT-g is an open-source KVM/Xen based full GPU virtualization solution; however, it lacks the feature of live migration. This work enables live migrating vGPUs for cloud applications such as virtual desktops, cloud gaming farms, cloud transcoding services, etc.
- Scaling up GPU Virtualization* [11, 8] (with Intel) Jan. 2016 – Mar. 2017
- This project scales up the maximum number of vGPUs in Intel GVT-g (i.e., Intel’s GPU virtualization technology) for 5× while minimizing the performance overhead. It makes GVT-g more consolidated, since more VMs with vGPU can be deployed on one physical machine.

EMPLOYMENT

- Intel Asia-Pacific Research & Development Ltd** *Shanghai, China*
Software Developer Intern July 2016 – June 2018
 Advisor: Dr. Yaozu Dong
- Intel Lab** *Hillsboro, Oregon*
Graduate Technical Intern June 2020 – Aug. 2020
 Advisor: Dr. Sanjay Kumar
- VMware Research** *Palo Alto, California*
Intern - Remote Memory May 2021 – Aug. 2021
 Advisors: Dr. Marcos K. Aguilera, Dr. Irina Calciu
- Alibaba DAMO Academy** *Sunnyvale, California*
Research Intern May 2022 – Aug. 2022
 Advisor: Dr. Dimin Niu

TEACHING

- Programming and Data Structure (SE 117)** *Shanghai, China*
Teaching Assistant Feb. 2016 – June 2016

RESEARCH MENTORING

Xiaohe Cheng (HKUST BSc → Google)	<i>2019</i>
Abel Mulugeta Eneyew (Addis Ababa Institute of Technology)	<i>2019</i>
Haoyang Zhang (UMich BSc → UIUC PhD)	<i>2021 – 2022</i>
Wentao Zhang (SJTU MSc)	<i>2022</i>
Yin Yuan (UMich BSc)	<i>2022</i>

PROFESSIONAL SERVICES

External Reviewer for International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS)	<i>2022</i>
Artifact Evaluation Committee Member for Journal of Systems Research (JSys)	<i>2021</i>
Artifact Evaluation Committee Member for Symposium on Operating Systems Principles (SOSP)	<i>2021</i>
Reviewer for IEEE Transactions on Parallel and Distributed Systems (TPDS)	<i>2018</i>

TECHNICAL SKILLS

Programming Language: C, C++, Verilog, System Verilog
OS & Virtualization: Linux Kernel Development, KVM, QEMU, Mediated Pass-Through
Program Analysis: LLVM, Klee, Yosys, Pyverilog