

HANGCHEN YU

<http://cs.utexas.edu/~hyu>
yuhc123@gmail.com

RESEARCH INTERESTS

Virtualization, Hypervisor, GPGPU/Accelerator virtualization, Operating systems

EDUCATION

The University of Texas at Austin, Austin, TX

Ph.D. student in Computer Science, UTCS Systems group

M.S. in Computer Science

Present

May 2020

Shanghai Jiao Tong University, Shanghai, China

B.S. in Information Engineering

June 2014

EXPERIENCE

Katana Graph Inc., Austin, TX

Engineering Internship

May 2020 - July 2020

- Built an AvA-based GPU service on AWS which consolidates GPU servers across Katana instances.

Microsoft Research, Redmond, WA

Research Internship

May 2018 - Aug. 2018

- Virtualized Intel Processor Trace for Hyper-V on SkyLake and post-SkyLake (e.g. IceLake) CPUs.
- Developed the *first* extensible processor supervisor feature state management (i.e. XSAVES support) for Hyper-V.
- The code has been checked-in to the production.

VMware Inc., Palo Alto, CA

Performance Engineering Internship

May 2017 - Aug. 2017

- Developed a log-independent tool to monitor and analyze the inter-service communications in vCenter on the fly.
- The tool requires no modifications to service codes, and leverages the generic information such as system calls.
- Visualized the results in the forms of responsive graphs and charts, and identified the optimization opportunities.

Google Inc., Madison, WI

Software Engineer Internship in Cloud Storage Team

May 2016 - Aug. 2016

- Designed, wrote and tested low-level modules in a distributed in-memory storage system (which serves AlphaGo) to detect and repair inconsistencies and corruptions in data blocks and locks.

GitCafe, D.G.Z Inc., Shanghai, China

Software Developer Internship

Dec. 2014 - June 2015

- Implemented concurrent pushes of repository to MoPaaS cloud platform and GitCafe pages service (gitcafe.io).
- Developed GitHub-compatible Markdown rendering engine, wikis/issues modules and reserved keyword manager.
- Developed task scheduler to achieve workload balance and live patching of git workers.

PUBLICATIONS

- [1] **H. C. Yu**, A. M. Peters, A. Akshintala and C. J. Rossbach, “**AvA: Accelerated Virtualization of Accelerators**,” to appear in *The 25th ACM International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS’20)*, March 2020.
- [2] A. Akshintala*, **H. C. Yu***, A. M. Peters, C. J. Rossbach, “**Trillium: The code is the IR**,” in *The Second Special Session on Virtualization in High Performance Computing and Simulation (HPCS VIRT’19)*, July 2019.
- [3] **H. C. Yu**, A. M. Peters, A. Akshintala and C. J. Rossbach, “**Automatic Virtualization of Accelerators**,” in *Proceedings of the Workshop on Hot Topics in Operating Systems (HotOS’19)*, no. 8, pp. 58-65, May 2019.
- [4] Y. J. Kwon, **H. C. Yu**, S. Peter, C. J. Rossbach and E. Witchel, “**Ingens: Huge Page Support for the OS and Hypervisor**,” in *SIGOPS Oper. Syst. Rev.*, vol. 51, no. 1, pp. 83-93, Aug. 2017.
- [5] **H. C. Yu**, and C. J. Rossbach, “**Full Virtualization for GPUs Reconsidered**,” in *14th Workshop on Duplicating, Deconstructing, and Debunking (WDDD)*, 2017.
- [6] Y. J. Kwon, **H. C. Yu**, S. Peter, C. J. Rossbach and E. Witchel, “**Coordinated and Efficient Huge Page Management with Ingens**,” in *USENIX Symposium on Operating Systems Design and Implementation (OSDI)*, 2016.

- [7] C. He, **H. C. Yu**, A. B. Li, X. L. Liang, J. P. Geng and R. H. Jin, “**Sideband Radiation Level Suppression in Time-Modulated Array by Nonuniform Period Modulation**,” in *IEEE Antennas Wirel. Propag. Lett.*, vol. 14, pp. 606-609, 2015.
- [8] L. Liu, R. H. Jin, **H. C. Yu**, X. Liang, J. Geng and X. Bai, “**A Compact Ultra-Wideband Power Divider with High Isolation**,” in *IEEE Antennas and Propagation Symposium*, Memphis, TN, USA, July 6-12 2014.
- [9] Y. X. Liu, C. L. Chen, **H. C. Yu** and X. P. Guan, “**Distortion Analysis for Delay Tolerant Data Collection for High-speed Wireless Sensor and Actor Networks**,” *Intelligent Control and Automation (WCICA), 2012 10th World Congress on*, pp. 4452-4457, July 2012.

SELECTED PROJECTS

AvA: Automatic Virtualization of Accelerators (HotOS’19, ASPLOS’20) *Sept. 2017 - Dec. 2019*

- Analyzed trends in accelerator and hypervisor evolution, observed important gaps in the virtualization design space.
- Enhanced API remoting with robust hypervisor-mediated virtualization to provide isolation and compatibility.
- Designed a system and a compiler for automatic construction of virtual accelerator stacks.
- Virtualized 10 accelerators and 12 APIs with the system, including GPU (CUDA, OpenCL), TPU (TensorFlow), a few edge devices, etc., while minimizing developer effort and maintaining near-native performance.

Multi-core Operating Systems Implementation *Aug. 2016 - Nov. 2016*

- Coded physical memory management based on Barrelfish kernel (a capability-based research OS by ETH Zürich).
- Spawned processes by memory allocator, capability and virtual address space initializer, ELF parser and dispatcher.
- Wrote RPC and memory server for centric memory allocation, and dynamic self-paging with page-fault handler.
- Booted multi-cores; established User-Level Message Passing and RPC between any two processes across the cores.
- Developed user-space network driver, ping server and UDP server to process and echo ICMP and UDP packets.

Ingens: Coordinated and Efficient Huge Page Management (OSDI’16) *Nov. 2015 - May 2016*

- Implemented Linux page walker to track the access frequency of pages and utilization of memory regions.
- Optimized performance of Java HotSpot 8 by enabling huge pages to different heap generations dynamically.
- Modified Redis to aggregate frequently accessed objects into continuous memory and promoted it into huge pages.

Toy Systems and Applications (Courses) *Oct. 2015 - Apr. 2016*

- Distributed systems: fault-tolerant chat room service (Paxos), weakly connected consistent key-value store (Bayou).
- Operating systems: lightweight x86 kernel, userspace networked file system (with FUSE), ELF program loader.
- Compilers: recursive descent parser (for Bali and SaM), x86-64 assembly code generator and C compiler.

TEACHING

| | |
|---|--------------------|
| <i>CS 314 Data Structures, Lead weekly discussion sections, UT Austin, Teaching Assistant</i> | <i>Fall 2017</i> |
| <i>CS 380L Advanced Operating Systems, UT Austin, Teaching Assistant</i> | <i>Fall 2017</i> |
| <i>CS 329E Data Visualization, UT Austin, Teaching Assistant</i> | <i>Fall 2016</i> |
| <i>CS 329E Elements of Programming Languages, UT Austin, Teaching Assistant</i> | <i>Spring 2016</i> |
| <i>CS 331 Algorithms and Complexity, UT Austin, Teaching Assistant</i> | <i>Spring 2016</i> |
| <i>CS 350C Advanced Computer Architecture, UT Austin, Teaching Assistant</i> | <i>Fall 2015</i> |

TALKS & POSTERS

| | |
|--|------------------|
| <i>(P) Automatic Accelerator Virtualization via Language and Compiler Support, Best Poster Award at IAP’19</i> | <i>Dec 2019</i> |
| <i>(T) Trillium: The code is the IR, HPCS VIRT’19</i> | <i>July 2019</i> |
| <i>(P) NightWatch: Remoting Accelerator APIs through the Hypervisor, OSDI’18</i> | <i>Oct 2018</i> |
| <i>(P) Efficient and Effective GPGPU Virtualization, UT Cloud Workshop’17</i> | <i>Nov 2017</i> |
| <i>(T) Full Virtualization for GPUs Reconsidered, WDDD ISCA’17</i> | <i>June 2017</i> |
| <i>(P) Ingens: Coordinated and Efficient Huge Page Management, OSDI’16</i> | <i>Nov 2016</i> |

AWARDS

| | |
|---------------------------|-------------|
| <i>OSDI Student Grant</i> | <i>2018</i> |
| <i>SOSP Scholarship</i> | <i>2017</i> |
| <i>OSDI Travel Grant</i> | <i>2016</i> |

SPECIALIZED SKILLS

| | |
|-------------------|--|
| Languages: | Spent most time with C; also proficient in C++, Python, Shell, Assembly, L ^A T _E X |
| Hobbies: | Exposure to Rust, Go, Java, Ruby Reading, tennis, table tennis, chess, badminton, basketball, DoTA |