Manager: Mike O'Connor

SUMMARY

Compute architecture researcher working on hardware-software co-design to improve system performance and energy efficiency.

EDUCATION

Ph.D. in Computer Science

Aug 2011 - Jul 2017

University of Pittsburgh, Pittsburgh, USA

• Thesis: "Addressing Prolonged Restore Challenges in Further Scaling DRAMs"

B.E. in Software Engineering

Sep 2007 - Jun 2011

Northwestern Polytechnical University (NPU), Xi'an, China

• Thesis: "DNA Cryptography based on DNA Fragment Assembly" (pub in ICIDT'2012)

EXPERIENCES

AMD® Inc [Work]

Seattle, USA

Postdoctoral Researcher, AMD Research

Aug 2017 - Present

- Optimize GPU architectures to improve performance and save energy for future Exascale computing

NVIDIA® Corporation [Intern]

Austin, USA

Research Intern, Architecture Research Group

May 2016 - Aug 2016

• Mentor: Niladrish Chatteriee

- Design memory system for next generation GPUs to achieve better latency tolerance

[GSR] University of Pittsburgh Pittsburgh, USA

Graduate Student Researcher, CS Department

May 2013 - Apr 2016

• Advisors: Prof. Youtao Zhang, Prof. Bruce Childers and Prof. Jun Yang (ECE)

- Improve system performance and energy efficiency via hardware-software co-designs

Alipay® Technology Inc., Alibaba [Intern]

Hangzhou, China

Java Developer Intern, Group Products Division

Aug 2010 – Dec 2010

- Implement a source management system based on SOFA/Spring framework

Optimize GPU Computation for Future Exascale Supercomputing

Projects [Exascale]

(Selected research projects during work and Phd study)

2017.08-

PathForward program of Department of Energy (DOE).

Design Memory System for Next Generation GPUs [GPU]

2016.05-2016.08

Envision the memory system requirements on future generation GPUs after Pascal; characterize and understand latency tolerance of GPU applications, and adapt cache designs to mitigate performance degradation of relaxed memory access latency. [Internal Reports & Presentations]

[Approx]

Apply Approximate Computing to Improve Memory Performance

2015.10-2016.08

Write Pintool to annotate variables in source codes and dynamically alter register and memory values to inject runtime errors; implement cache and virtual memory to collect memory access traces, and adapt memory simulator to report performance and energy results. [PACT'2017, MemSys'2016]

[Memory] Improve performance and energy in DRAM and NVM

2013.2-2015.9

Perform pioneering studies on DRAM further scaling issues via modeling and simulation [HPCA'2016, TO-DAES'2017, DATE'2015]; propose encodings to shorten PCM write latency, and re-organize Domain Wall Memory (DWM) to reduce cache access energy [ISLPED'2013, ICCD'2015].

Research

Memory System, GPU, Computer Architecture and Systems, Software-Hardware Co-design

Publications

[C9]

9 conference, 1 journal and 1 poster papers (full-list, Google Citation, DBLP)

HPCA'2018

Anthony Gutierrez, Bradford Beckmann, ...(11), Xianwei Zhang, Matt Sinclair - Lost in Abstraction: Pitfalls of Analyzing GPUs at the Intermediate Language Level. The 24th IEEE International Symposium on High-Performance Computer Architecture, Vienna, Austria, 2018.

Xianwei Zhang, Youtao Zhang, Bruce R. Childers and Jun Yang [C8]

PACT'2017

- DrMP: Mixed Precision-Guided DRAM Restore for High Performance Approximate

and Precise Computing. The 26th International Conference on Parallel Architectures and Compilation Techniques, Portland, Oregon, USA, 2017.

Xianwei Zhang, Curriculum Vitae, p. 1/2

- [J1] Xianwei Zhang, Youtao Zhang, Bruce R. Childers and Jun Yang TODAES'2017
 On the Restore Time Variations of Future DRAM Memory. ACM Trans. on Design Automation of Electronic Systems, Vol. 22(2), 26:1-26:24.
- [C7] Xianwei Zhang, Youtao Zhang, Bruce R. Childers and Jun Yang
 Restore Truncation for Performance Improvement in Future DRAM Systems. The 22nd IEEE Symp. on High Performance Computer Architecture, Barcelona, Spain, 2016.
- [C6] Xianwei Zhang, Youtao Zhang, Bruce R. Childers and Jun Yang
 AWARD: Approximation-aWAre Restore in Further Scaling DRAM. The International Symposium on Memory Systems, Washington D.C., USA, October 2016.
- [C5] Xianwei Zhang, Youtao Zhang, Bruce R. Childers and Jun Yang
 Exploiting DRAM Restore Time Variations in Deep Sub-micron Scaling. The IEEE conference on Design, Automation and Test in Europe, Grenoble, France, 2015.
- [C4] Xianwei Zhang, Youtao Zhang and Jun Yang
 DLB: Dynamic Lane Borrowing for Improving Bandwidth and Performance in Hybrid Memory
 Cube. The 33rd IEEE Int'l Conf. on Computer Design, New York City, USA, 2015.
- [C3] Xianwei Zhang, Youtao Zhang and Jun Yang
 TriState-SET: Proactive SET for Improved Performance in MLC Phase Change Memories. The 33rd IEEE Int'l Conf. on Computer Design, New York City, USA, 2015.
- [C2] Xianwei Zhang, Youtao Zhang and Jun Yang
 Exploit Common Source-Line to Construct Energy Efficient Domain Wall Memory based
 Caches. The 33rd IEEE Int'l Conf. on Computer Design, New York City, USA, 2015.
- [W1] Xianwei Zhang, Youtao Zhang and Jun Yang
 Adaptive Lane Borrowing of Hybrid Memory Cube. The 52nd ACM/IEEE Design Automation Conference (DAC), San Francisco, USA, 2015.
- [C1] Xianwei Zhang, Lei Jiang, Youtao Zhang, Chuanjun Zhang and Jun Yang ISLPED'2013
 WoM-SET: Lowering Write Power of Proactive-SET based PCM Write Strategy Using WoM
 Code. The 19th Int'l Symp. on Low Power Electronics and Design, Beijing, China, 2013.

 *** Best Paper Award ***

COMP SKILLS Programming: C/C++, JAVA, Linux/Shell/AWK, Python, Android, SQL, R

Tools: Makefile, gcc/g++, GDB, Varius, Vim, Git, gem5, Intel Pin, IATEX

Artifacts: Framework of DRAM scaling study, Pin tool of assembly operation

Real-time Twitter posts using Arduino and sensors

Python pub-quality plotting tool, Motion-based Android App, etc.

Honors & Andrew Mellon Predoctoral Fellowship

AWARDS

Misc

- awarded to Phd students of exceptional achievement and promise

Student Travel Awards *HPCA*'2016, *SPAA*'2015, *CS Dept.*'2016&2015

Best Paper Award ISLPED'2013

- based on the rating of anonymous reviewers and a panel of judges

Recipient of 2011 graduation design (Thesis) key support fund

- small research grant for undergraduate thesis project, 2.5% funding rate

Tencent® Technology Excellence Scholarship - top grade, 3 winners NPU-wide

Homepage: https://xianweiz.github.io

Github: https://github.com/xianweiz

Blog: http://iarchsys.com

Linkedin: https://www.linkedin.com/in/xianweizhang/

University of Pittsburgh'2016

NPU'2011

Tencent Inc.'2009