

# ZHENGRONG WANG

404 Westwood Plaza, EVI 468 – 90095, Los Angeles, CA, USA  
seanzw@ucla.edu <https://seanzw.github.io> Google Scholar

## EDUCATION

---

**University of California, Los Angeles, Department of Computer Science** Los Angeles, USA  
Ph.D. in Computer Science, Advisor: Tony Nowatzki Aug. 2018 - Nov. 2023  
*Dissertation: General, Flexible and Unified Near-Data Computing*

**University of California, Los Angeles, Department of Computer Science** Los Angeles, USA  
Master of Science in Computer Science Sep. 2016 - Jul. 2018  
*Thesis: An LLVM-IR Datagraph-Based Simulator for Flexible Design Space Exploration over Accelerator Architectures*

**Tsinghua University, Department of Electronic Engineering** Beijing, China  
Bachelor of Engineering in Electronic Engineering, GPA: 91/100 Aug. 2012 - Jul. 2016  
*Thesis: Optimizing Convolutional Neural Network on FPGA under Heterogeneous Computing Framework with OpenCL*

## PROFESSIONAL SKILLS

---

Mathematic: Familiar with calculus, linear algebra, probability theory, discrete mathematics, algorithms.  
Research Areas: Computer architecture, compiler, cache, microarchitecture, network-on-chip, CPU, GPU.  
Programming: Skilled at C/C++, Python, assembly, LLVM IR.

## PROFESSIONAL EXPERIENCES


---

**Nvidia Research** Jun. 2022 - Sep. 2022  
Research Scientist, Mentor: Neal Crago, Manager: Steve Keckler


- Examine memory bottleneck in GPU for key machine learning kernels.
- Build a prototype of an enhanced tensor memory accelerator (TMA).
- Evaluted with state-of-the-art point cloud applications.

## OPEN SOURCE PROJECTS & INFRASTRUCTURES

---

**Gem5-AVX**  Jan. 2019 - Present  
First Author & Maintainer

- Add AVX-512 support to gem5 simulator, extensively used in research.
- Faithfully model the microarchiecture of vectorized instructions, including microops.
- Detailed tutorials on how to support new instructions.

**Stream-Specialized Near-Data Acceleration Framework**  Jan. 2018 - Present  
First Author & Maintainer

- Full-stack implementation of stream-specialized near-data acceleration.
- Include LLVM-based compiler transformation and end-to-end simulation in gem5.
- Results published in ISCA' 19, HPCA' 21, HPCA' 22, MICRO '23, ASPLOS '23.

## PUBLICATION

---

Affinity Alloc: Taming Not-So Near-Data Computing  
**Zhengrong Wang**, Christopher Liu, Nathan Beckmann, Tony Nowatzki  
*IEEE/ACM International Symposium on Microarchitecture (MICRO)*, 2023, Toronto, Canada.

Infinity Stream: Portable and Programmer-Friendly In-/Near-Memory Fusion  
**Zhengrong Wang**, Christopher Liu, Aman Arora, Lizy John, Tony Nowatzki  
*ACM International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS)*, 2023, Vancouver, Canada.

Infinity Stream: Enabling Transparent and Automated In-Memory Computing  
**Zhengrong Wang**, Christopher Liu, Tony Nowatzki  
*IEEE Computer Architecture Letters*, Vol. 21, No. 2, 2022.

OverGen: Improving FPGA Usability through Domain-specific Overlay Generation  
 Sihao Liu, Jian Weng, Dylan Kupsh, Atefeh Sohrabizadeh, **Zhengrong Wang**, Licheng Guo, Jiuyang Liu, Maxim Zhulin, Lucheng Zhang, Jason Cong, Tony Nowatzki  
*IEEE/ACM International Symposium on Microarchitecture (MICRO)*, 2022, Chicago, USA.

**Best Paper Runner-Up**

Near-Stream Computing: General and Transparent Near-Cache Acceleration  
**Zhengrong Wang**, Jian Weng, Sihao Liu, Tony Nowatzki  
*IEEE International Symposium on High-Performance Computer Architecture (HPCA)*, 2022, Seoul, South Korea.

Stream Floating: Enabling Proactive and Decentralized Cache Optimizations  
**Zhengrong Wang**, Jian Weng, Jason Lowe-Power, Jayesh Gaur, Tony Nowatzki  
*IEEE International Symposium on High-Performance Computer Architecture (HPCA)*, 2021, Seoul, South Korea.

**Best Paper Runner-Up**

DSAGEN: Synthesizing Programmable Spatial Accelerators  
 Jian Weng, Sihao Liu, Vidushi Dadu, **Zhengrong Wang**, Preyas Shah, Tony Nowatzki  
*ACM International Symposium on Computer Architecture (ISCA)*, 2020, virtual.

**IEEE Micro Top Picks Honorable Mention**

A Hybrid Systolic-Dataflow Architecture for Inductive Matrix Algorithms  
 Jian Weng, Sihao Liu, **Zhengrong Wang**, Vidushi Dadu, Tony Nowatzki  
*IEEE International Symposium on High-Performance Computer Architecture (HPCA)*, 2020, San Diego, USA.

Stream-Based Memory Access Specialization for General Purpose Processors  
**Zhengrong Wang**, Tony Nowatzki  
*ACM International Symposium on Computer Architecture (ISCA)*, 2019, Phoenix, USA.

The Gem5 Simulator: Version 20.0+  
 Jason Lowe-Power, Abdul Mutaal Ahmad, Ayaz Akram, ..., **Zhengrong Wang**, et al.  
*arXiv:2007.03152v2*, 2020.

Optimizing Convolutional Neural Network on FPGA under Heterogeneous Computing Framework with OpenCL  
**Zhengrong Wang**, Fei Qiao, Zhen Liu, Yuxiang Shan, Xunyi Zhou, Li Luo, Huazhong Yang  
*IEEE Region 10 Conference (TENCON)*, 2016, Singapore.

## AWARDS AND HONORS

---

Dissertation Year Fellowship, <i>UCLA</i>	June. 2023
Best Paper Runner-Up (OverGen, in <i>MICRO '22</i> ), <i>IEEE</i>	Oct. 2022
Best Paper Runner-Up (Stream Floating, in <i>HPCA '21</i> ), <i>IEEE</i>	Feb. 2021
IEEE Micro Top Picks 2020 Honorable Mention (DSAGEN, in <i>ISCA '20</i> ), <i>IEEE</i>	Jan. 2021
2021 Dongguan Entrepreneur Scholarship, <i>Dongguan Entrepreneurs Federation</i>	Nov. 2021
Second-class Scholarship for Excellent Freshmen, <i>Tsinghua University</i>	Oct. 2012
Wang Zhaosheng Scholarship for Excellent Student from Dongguan, <i>Wang Zhaosheng Foundation</i>	Oct. 2012