

Yueming Hao

🏠 about.findhao.net | ✉ yhao24@ncsu.edu

Research Interests

- Performance Analysis and Optimizations for GPU Applications
- Data-Centric Inefficiencies Analysis
- High Performance Computing

Education

North Carolina State University

Expected Degree: Ph.D. in Computer Science;

Advisors: Prof. Xu Liu

Raleigh, North Carolina, USA

Aug. 2020 - Present

College of William and Mary

Ph.D. Study in Computer Science;

Advisors: Prof. Xu Liu

Williamsburg, Virginia, USA

Aug. 2019 - Aug.2020

Shandong University

M.E. in Computer Science and Technology; Advisors: Prof. Lei Ju

Jinan, Shandong, China

Sep. 2016 - June. 2019

Shandong University

B.E. in Computer Science and Technology; Advisors: Prof. Lei Ju

Jinan, Shandong, China

Sep. 2012 - June. 2016

Research Experience

North Carolina State University

August 2020 - Present

- **VALUEXPRT: Exploring Value Patterns in GPU-accelerated Applications**
(submitted to ASPLOS'2022)
 - Categorized eight value patterns in GPU-accelerated Applications
 - Proposed a new profiling tool to analyze value patterns and value flows to pinpoint value-related inefficiencies.
 - Achieved non-trivial speedups and upstream our optimizations to benefit the communities, like **PyTorch**, **darknet**, etc.
- **GPUPUNK: A Unified Memory Page False Sharing Profiler for CPU-GPU Platforms**
(working on it now)
 - Combined CPU and GPU instrumentation techniques and analyzes the memory accesses crossing CPU and GPU

Meta, Student Researcher

September 2022 - December 2022

Meta, Research Scientist Intern

May 2022 - August 2022

- Maximizing GPU Utilization in Deep Learning Frameworks

(target ASPLOS 2023)

- Did a characteristic study for machine learning models
- Developed a PyTorch profiler

NVIDIA, Research Intern

May 2020 - August 2020

- **DRGPU: A Top-Down Profiler for GPU**

- Quantified stall cycles and decomposes them according to various hardware events for root causes.
- Provided focused, hierarchical performance deficit attribution with minimum manual interference.

College of William and Mary

August 2019 - May 2020

- **GVPROF: A Value Profiler for GPU-based Clusters**

(SC'2020)

- Systematically studied temporal and spatial value redundancies in GPU codes for both memory loads/stores and proposed various techniques for optimization.
- Proposed GVProf, the first value profiler for NVIDIA GPUs.
- Designed GVProf to provide useful performance insights, including derived redundancy metrics, full calling contexts, and a data-centric view for instructions and data objects.

Publication

ASPLOS'2022 "ValueExpert: Exploring Value Patterns in GPU-accelerated Applications", ASPLOS22: Architectural Support for Programming Languages and Operating Systems, 2022.
Keren Zhou, **Yueming Hao***, John Mellor-Crummey, Xiaozhu Meng, Xu Liu. (*co-first authors)

SC'2020 "GVProf: A value profiler for GPU-based clusters." *SC20: International Conference for High Performance Computing, Networking, Storage and Analysis*, 2020.
Keren Zhou; **Yueming Hao**; John Mellor-Crummey; Xiaozhu Meng; Xu Liu

Honors & Awards

2022 Distinguished Artifact Award, ASPLOS 2022
2021 Runner Up, A-HUG Cloud HPC Hackathon
2021 NCSU Summer Graduate Merit Award (GMA)
2014 First Prize of China Undergraduate Mathematical Contest in Modeling

Skills

Programming Languages: C, C++, Python
HPC Programming Models: CUDA, OpenMP, MPI