

# Keren Zhou

6100 Main ST – Houston, TX – 77005, United States

☎ +1-281-687-6961

✉ kerezhou@outlook.com

🌐 www.jokeren.tech

## EDUCATION

09/2017-07/2023	Department of Computer Science, Rice University	Houston, United States
	Expected Degree: <i>Ph.D. in Computer Science</i>	Advisor: John Mellor-Crummey
09/2014-07/2017	Institute of Computing Technology, Chinese Academy of Sciences	Beijing, China
	Degree: <i>M.S. in Computer Architecture</i>	Advisor: Guangming Tan
09/2010-07/2014	School of Software, Yunnan University	Kunming, China
	Degree: <i>B.E. in Network Engineering</i>	Advisor: Wei Zhou

## RESEARCH EXPERIENCE

09/2017-NOW	Rice University	Houston, United States
	<b>A Performance Analysis Tool for GPU-accelerated Supercomputers</b>	
	<ul style="list-style-type: none"><li>Extended HPCToolkit to support measurement and analysis of accelerated OpenMP and CUDA GPU programming models in a large-scale heterogeneous environment;</li><li>Built a profile view of GPU program executions and attributed runtime samples to the corresponding calling context.</li></ul>	
06/2015-07/2017	Nvidia-Sugon-ICT Deep Learning Joint Laboratory	Beijing, China
	<b>Deep Learning Accelerating Packages</b>	
	<ul style="list-style-type: none"><li>Built a performance analysis model to estimate GPU kernels' performance bottlenecks;</li><li>Devised fine-grained vectorization and blocking on GPUs and CPUs to accelerate CNNs.</li></ul>	

## INDUSTRY EXPERIENCE

06/2018-08/2018	Facebook Inc.	Menlo Park, United States
	<ul style="list-style-type: none"><li>Accelerated neural networks on ARM CPUs using auto-tuning methods;</li><li>Reference: Research Scientist Hao Lu, hlu@fb.com.</li></ul>	
04/2017-07/2017	Nvidia Inc.	Beijing, China
	<ul style="list-style-type: none"><li>Developed quantization tools on emerging GPUs to utilize INT8 capabilities;</li><li>Reference: Technical Manager Julien Lai, julienlai@nvidia.com.</li></ul>	
10/2013-02/2014	Baidu Inc.	Beijing, China
	<ul style="list-style-type: none"><li>Optimized Hadoop workflow and improved its performance by 30%;</li><li>Reference: Senior Engineer Jing Li, lijing16@baidu.com.</li></ul>	

## SELECTED PUBLICATIONS

- [1] **Keren, Zhou**; Guangming, Tan; Wei, Zhou: Quadboost: A Scalable Concurrent Quadtree. In: *IEEE Transactions on Parallel and Distributed Systems* (TPDS), 2018
- [2] **Keren Zhou**; Guangming Tan; Xiuxia Zhang; Chaowei Wang; Ninghui Sun: A Performance Analysis Framework for Exploiting GPU Microarchitectural Capability. In *26th ACM International Conference on Supercomputing* (ICS), 2017
- [3] Xiuxia, Zhang; Guangming, Tan; Shuangbai, Xue; Jiajia, Li; **Keren, Zhou**; Mingyu, Chen: Understanding GPU Microarchitecture to Achieve Bare-Metal Performance Tuning. In: *22nd ACM SIGPLAN Symposium on Principles and Practice of Parallel Programming* (PPOPP), 2017

## AWARDS & HONORS

2019	Second Place, ACM CGO Student Research Competition
2017	Ken Kennedy Institute Andrew Ladd Fellowship
2017	Ken Kennedy Institute CS&E Fellowship
2011&2012&2016	National Scholarship
2016	Schlumberger Scholarship
2014	Outstanding B.E. Degree Thesis of Yunnan University
2013	Meritorious Winner, Mathematical Contest in Modeling
2011&2012	Merit Student of Yunnan Province