Mingzhe Zhang

Ph.D, Associate Professor

No.65 Xingshikou Road, Haidian District, Beijing, China
100190
✉ *zhangmingzhe@iie.ac.cn*
Institute of Information Engineering
Chinese Academy of Science

Education

2013–2018	**Ph.D of Computer Science**, *Institute of Computing Technology, Chinese Academy of Science (ICT,CAS)*, Beijing, China.
2009–2013	**Master of Computer Science**, *Inner Mongolian University*, Hohhot, China.
2004–2008	**Bachelor of Computer Science**, *Nanjing University of Posts and Telecommunications*, Nanjing, China.

Professional Experience

2021.4–present	**Associate Professor**, *Institute of Information Engineering, Chinese Academy of Sciences (IIE, CAS)*.
2018.7–2021.4	**Assistant Professor**, *Institute of Computing Technology, Chinese Academy of Sciences (ICT, CAS)*.
2015.11–2017.2	**Visiting Scholarship**, *Department of Computer Science, University of Chicago*.
2010.10–2013.6	**Research Intern**, *Institute of Computing Technology, Chinese Academy of Science (ICT,CAS)*.
2008.7–2009.9	**Engineer**, *Institute of Software, Chinese Academy of Science (ISCAS)*.

Honors and Awards

2020	**Alibaba-CCF TCIST Outstanding Paper Award.**.
2019	**Runner-up of CCF Outstanding PhD Thesis Award. (The only selected candidate for computer architecture related.)**.
2017	**National Scholarship for Outstanding PhD Student**.
2017	**Outstanding Student of University of Chinese Academy of Sciences**.
2015	**National Scholarship for Studying Abroad**.
2015	**Outstanding Student of University of Chinese Academy of Sciences**.
2012	**Excellent Presentation Award of 2nd Ph.D Forum of Micro-electronic in Beijing**.
2012	**Outstanding Student of State Key Laboratory of Computer Architecture, ICT, CAS**.
2011	**Outstanding Student of State Key Laboratory of Computer Architecture, ICT, CAS**.

Research Interests

Computer Architecture	Micro-Architecture, NoC, Cache, Memory System, Processing-In-Memory, ORAM
Non-Volatile Memory	Phase Change Memory (PCM), Memristor, STT-RAM
System & Accelerator	CGRA, Mapping Scheme, Accelerator for Bioinformatics

Students

Current:

- Shengyu Fan, M.S. student, co-advise with Prof. Weizhi Xu since Summer 2021.
- Yu Deng, M.S. student, since Autumn 2021.

Former:

- Mo Zou, Ph.D student, co-advise with Prof. Zhimin Tang from Fall 2019 to Spring 2021.
- Long Tan, Ph.D student, co-advise with Prof. Dongrui Fan from Fall 2020 to Spring 2021.
- Ziqi Wang, M.S. student, co-advise with Prof. Hao Zhang from Spring 2020 to Spring 2021.
- Dingyuan Cao, undergraduate research intern from Tsinghua University, from Spring 2019 to Winter 2020. (Current: Ph.D student of UIUC, under the supervision of Prof. Josep Torrellas.)

Grants

2021–2013 **Research on Phase-Chang Memory Optimization Technology for High-Throughput Application Scenarios**, *240000 CNY*, PI.
NSFC (Grant No. 62002339)

2020–2014 **System Software Stack and Hardware Operator for Graph Computing based on Processing-In-Memory Architecture**, *4680000 CNY*, PI.
Sub Project of Strategic Priority Research Program (Type-B), Chinese Academy of Sciences: Frontier Sciences in Fundamental Devices and Systems for In-Memory Computing (XDB44000000)

2020–2021 **Optimization for Non-Volatile Memory based on Device Characteristics**, *50000 CNY*, PI.
Innovation Project of State Key Laboratory of Computer Architecture

2018–2019 **Acceleration for Graph Computing based on Processing-In-Memory Architecture**, *400000 CNY*, Co-PI (PI: Prof. Xiaochun Ye).
Sub Project of Strategic Priority Research Program (Incubation Project), Chinese Academy of Sciences: The Frontier Science and Technology for Storage-Driven Computing (XDPB12)

Academia Activities

- **[Invited Presentation]** FindeR: Accelerating FM-Index-based Exact Pattern Matching in Genomic Sequences through ReRAM technology. The 25th National Conference of Information Storage (NCIS2019). Shenzhen, China. 2019.
- **[Conference Presentation]** Balancing Performance and Lifetime of MLC PCM by Using a Region Retention Monitor. HPCA 2017. Austin, TX. 2017.

- **[Conference Presentation]** A Path-Adaptive Opto-electronic Hybrid NoC for Chip Multi-processor. ISPA 2013. Melbourne, Australia. 2013.
- **[Conference Presentation]** Self-Correction Trace Model: A Full-System Simulator for Optical Network-on-Chip. IPDPSW 2012. Shanghai, China. 2012.

Services

Conference Organization Committee

- Session Chair, The 17th Annual IFIP International Conference on Network and Parallel Computing (NPC), 2020
- Session Chair, International Conference on Networking, Architecture, and Storage (NAS), 2019
- Session Chair, IEEE Non-Volatile Memory Systems and Applications Symposium (NVMSA), 2019

Conference Technical Program Committee

- IEEE International Conference on Parallel and Distributed Systems (ICPADS), 2021
- IEEE Non-Volatile Memory Systems and Applications Symposium (NVMSA), 2021
- The 39th IEEE International Conference on Computer Design (ICCD 2021), 2021
- IEEE Non-Volatile Memory Systems and Applications Symposium (NVMSA), 2020
- The 17th Annual IFIP International Conference on Network and Parallel Computing (NPC), 2020
- International Conference on Networking, Architecture, and Storage (NAS), 2019
- IEEE Non-Volatile Memory Systems and Applications Symposium (NVMSA), 2019

Journal Reviewer

- IEEE Transactions on Computers (TOC)
- IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (TCAD)
- IEEE Transactions on VLSI (TVLSI)
- ACM Transactions on Embedded Computing Systems (TECS)
- IEEE Micro
- Journal of Computer Science and Technology (JCST)
- Journal of Communications and Networks (JCN)
- Journal of Computer Research and Development (CRAD)

Publications

* *Highlights: totally 25 paper, 2 HPCA, 1 MICRO, 2 PACT, 3 ICCD, 1 TOC, 1 TCAD, 1 TPDS.*
- **[MICRO 2021]** Hang Lu, Liang Chang, Chenglong Li, Zixuan Zhu, Yanhuan Liu, Shengjian Lu, **Mingzhe Zhang**. "Distilling Bit-level Sparsity Parallelism for General Purpose Deep Learning Acceleration", 54th IEEE/ACM International Symposium on Microarchitecture. 2021.
- **[ICPP 2021]** Hongyan Li, Hang Lu, Jiawen Huang, Wenxu Wang, **Mingzhe Zhang**, Wei Chen, Liang Chang, Xiaowei Li. "BitX: Empower Versatile Inference with Hardware Runtime Pruning", 50th International Conference on Parallel Processing. 2021.

- **[ISLPED 2021]** Liang Yan, **Mingzhe Zhang**, Rujia Wang, Xiaoming Chen, Xingqi Zou, Xiaoyang Lu, Yinhe Han, and Xian-He Sun. "CoPIM: A Concurrency-aware PIM Workload Offloading Architecture for Graph Applications", International Symposium on Low Power Electronics and Design. 2021.
- **[HPCA 2021]** Dingyuan Cao, **Mingzhe Zhang**, Hang Lu, Xiaochun Ye, Dongrui Fan, Yuezhi Che, and Rujia Wang. "Streamline Ring ORAM Accesses through Spatial and Temporal Optimization", IEEE International Symposium on High Performance Computer Architecture. 2021.
- **[arXiv]** Taoran Xiang, Lunkai Zhang, Shuqian An, Xiaochun Ye, **Mingzhe Zhang**, Yanhuan Liu, Mingyu Yan et al. "RISC-NN: Use RISC, NOT CISC as Neural Network Hardware Infrastructure." arXiv preprint arXiv:2103.12393 (2021).
- **[IEEE TCAD]** Hang Lu, **Mingzhe Zhang(*)**, Xin Wei, Guihai Yan, Qi Wang, Huawei Li and Xiaowei Li. "Architecting Effectual Computation for Machine Learning Accelerators", IEEE Transactions on Computer-Aided Design.(accepted). 2020.
- **[ISPA 2019]** Chundian Li, **Mingzhe Zhang**, Zhiwei Xu, Xianhe Sun. "Self-adaptive Address Mapping Mechanism for Access Pattern Awareness on DRAM", 17th IEEE International Symposium on Parallel and Distributed Processing with Applications. 2019.
- **[ICCD 2019]** **Mingzhe Zhang**, Lunkai Zhang, Frederic T. Chong, Zhiyong Liu. "Balancing Performance and Energy Efficiency of ONoC by Using Adaptive Bandwidth", 37th IEEE International Conference on Computer Design. 2019.
- **[ICCD 2019]** Ning Lin, Hang Lu, Jingliang Gao, **Mingzhe Zhang**, Xiaowei Li. "When Deep Learning Meets the Edge: Auto-Masking Deep Neural Networks for Efficient Machine Learning on Edge Devices", 37th IEEE International Conference on Computer Design. 2019.
- **[HPCC 2019]** Shuqian An, **Mingzhe Zhang (co-first author)**, Xiaochun Ye, Da Wang, Hao Zhang, Dongrui Fan, Zhimin Tang. "C-MAP: Improving the Effectiveness of Mapping Method for CGRA by Reducing NoC Congestion", 21st IEEE International Conference on High Performance Computing and Communications. 2019.
- **[PACT 2019]** Farzaneh Zokaee, **Mingzhe Zhang**, Lei Jiang. "FindeR: Accelerating FM-Index-based Exact Pattern Matching in Genomic Sequences through ReRAM technology", The 28th International Conference on Parallel Architectures and Compilation Techniques. 2019.
- **[CRAD]** **Mingzhe Zhang**, Fa Zhang, Zhiyong Liu. "A Survey on Architecture Research of Non-Volatile Memory based on Dynamical Trade-off" (in Chinese), Journal of Computer Research and Development. 2019.
- **[IEEE TOC]** **Mingzhe Zhang**, Lunkai Zhang, Lei Jiang, Frederic T. Chong, Zhiyong Liu. "Quick-and-Dirty: An Architecture for High-Performance Temporary Short Writes in MLC PCM", IEEE Transactions on Computers. 2019.
- **[DAC 2019]** Farzaneh Zokaee, **Mingzhe Zhang**, Xiaochun Ye, Dongrui Fan, Lei Jiang. "Magma: A Monolithic 3D Vertical Heterogeneous ReRAM-based Main Memory Architecture", Proceedings of the 56th Annual Design Automation Conference. 2019.
- **[BIBM 2018]** Zihao Wang, **Mingzhe Zhang**, Jingrong Zhang, Rui Yan, Xiaohua Wan, Zhiyong Liu, Fa Zhang, Xuefeng Cui. "Mmalloc: A Dynamic Memory Management on Many-core Coprocessor for the Acceleration of Storage-intensive Bioinformatics Application", IEEE International Conference on Bioinformatics and Biomedicine". 2018.
- **[ICCD 2017]** **Mingzhe Zhang**, Lunkai Zhang, Lei Jiang, Frederic T. Chong, Zhiyong Liu, "Quick-and-Dirty: Improving Performance of MLC PCM by Using Temporary Short Writes", 35th IEEE International Conference on Computer Design. 2017

- **[HPCA 2017]** <u>**Mingzhe Zhang**</u>, Lunkai Zhang, Lei Jiang, Zhiyong Liu, Frederic T. Chong, "Balancing Performance and Lifetime of MLC PCM by Using a Region Retention Monitor", IEEE International Symposium on High Performance Computer Architecture. 2017.
- **[IGSC 2016]** <u>**Mingzhe Zhang**</u>, Yangguang Shi, Fa Zhang, Zhiyong Liu, "COM-RANCE: A rapid method for Network-on-Chip design space exploration", 7th International Green and Sustainable Computing Conference. 2016.
- **[IEEE TPDS]** Shaoli Liu, Tianshi Chen, Ling Li, Xi Li, <u>**Mingzhe Zhang**</u>, Chao Wang, Haibo Meng, Xuehai Zhou, and Yunji Chen, "FreeRider: Non-local Adaptive Network-on-Chip Routing with Packet-Carried Propagation of Congestion Information", IEEE Transaction on Parallel and Distributed Systems. 2015.
- **[PACT 2014]** Lunkai Zhang, Dmitri B. Strukov, Hebatallah Saadeldeen, Dongrui Fan, **Mingzhe Zhang**, Diana Franklin, "SpongeDirectory: flexible sparse directories utilizing multi-level memristors", 23rd International Conference on Parallel Architecture and Compilation Techniques. 2014.
- **[ISLPED 2013]** Xiaochun Ye, Dongrui Fan, Ninghui Sun, Shibin Tang, **Mingzhe Zhang**, Hao Zhang, "SimICT: A fast and flexible framework for performance and power evaluation of large-scale architecture". International Symposium on Low Power Electronics and Design. 2013.
- **[NAS 2013]** Lunkai Zhang, **Mingzhe Zhang**, Lingjun Fan, Da Wang, Paolo Ienne, "Spontaneous Reload Cache: Mimicking a Larger Cache with Minimal Hardware Requirement". IEEE 8th International Conference on Networking, Architecture and Storage. 2013.
- **[ISPA 2013]** Shuai Zhang, Zhiyong Liu, Dongrui Fan, Fenglong Song, **Mingzhe Zhang**, "Energy-Performance Modeling and Optimization of Parallel Computing in On-Chip Networks". IEEE International Symposium on Parallel and Distributed Processing with Applications. 2013.
- **[ISPA 2013]Mingzhe Zhang**, Da Wang, Xiaochun Ye, Liqiang He, Dongrui Fan, Zhiyong Liu, "A Path-Adaptive Opto-electronic Hybrid NoC for Chip Multi-processor". IEEE International Symposium on Parallel and Distributed Processing with Applications. 2013.
- **[IPDPSW 2012]Mingzhe Zhang**, Liqiang He, Dongrui Fan, "Self-Correction Trace Model: A Full-System Simulator for Optical Network-on-Chip". 19th Reconfigurable Architectures Workshop (collocated with IPDPS 2012). 2012.