Jie Zhao

Curriculum Vitae

State Key Laboratory of Mathematical Engineering and Advanced Computing,
No. 62, Kexuedadao Street
450001, Zhengzhou, China
✉ *yaozhujiajie@gmail.com*
https://yaozhujia.github.io/

Empolyments

2016.06–present **Assitant Professor**, *State Key Laboratory of Mathematical Engineering and Advanced Computing*, Zhengzhou.

Research Interests

Dr. Jie Zhao is an Assistant Professor at the State Key Laboratory of Mathematical Engineering and Advanced Computing (SKL-MEAC), Zhengzhou, China. He has been working on compilation techniques since he was a graduated student. He devotes to high-performance code generation for different architectures using compilation approaches and strives to bridge the gap between high-level programming models and underlying computer architectures. Specifically, his research interests and experiences span across:

- code generation and optimization
- deep learning systems
- numerical program analysis

Grants

NSFC 2021 **Principle Investigator of the SKL-MEAC part**, *Deep Learning and Tensor Compilers based on the Polyheral Model*, National Natural Science Foundation of China, No. U20A20226, 2021.01-2024.12.
2,600,000 CNY in total for the joint (with Tsinghua University and Beijing OneFlow Research.) project, and 800,000 CNY for the SKL-MEAC part

NSFC 2018 **Researcher**, *Analysis and Optimization of the Precision of Mathematical Functions on Domestic Processors*, National Natural Science Foundation of China, No. 61802434, 2019.01-2021.12.
250,000 CNY

NSFC 2017 **Principle Investigator**, *Polyhedral Compilation Techniques for Heterogeneous Architectures*, National Natural Science Foundation of China, No. 61702546, 2018.01-2020.12.
240,000 CNY

Experience

2020.04–present **Senior Consultant**, *Stream Computing Co., Ltd.*, Hangzhou, China.

2019.02–2021.12 **Senior Consultant**, *Huawei Technologies Co., Ltd.*, Hangzhou, China.

2020.04–2021.04 **Visiting Scholar**, *Alibaba Group*, Beijing, China.

2018.05–2018.09 **Consultant**, *Huawei Technologies France SASU*, Paris, France.

Education

2015.03–2018.12 **PhD**, *INRIA & École Normale Supérieure*, 75005 Pairs, France, PhD Thesis: A Combined Language and Polyhedral Approach for Heterogeneous Parallelism.
Supervised by Prof. Albert Cohen

2009.09–2012.06 **MSc**, *National Digital Switching System Engineering & Technological Research Center (NDSC)*, 450001, Zhengzhou, China, Master thesis: Research on Optimization Technologies of Parallel Compilation for Distributed Memory Architecture.
Supervised by Prof. Rongcai Zhao

2005.08–2009.07 **BSc**, *Department of Computer Science and Technology, Tsinghua University*, 100084, Beijing, China.

Awards

2020 IEEE/ACM MICRO-53 Best Paper Nominees

2019 HPC China 2019 Outerstanding Paper Award

2017 SIGPLAN Grant for PLDI 2017

2016 SIGPLAN PLMW Scholarship for OOPSLA 2016

2014 NDSC Scholarship

2013 NDSC's Excellent Master Dissertation Award, 1st place

Languages

Mandarin	**Mothertongue**	*Daily use*
English	**Flent**	*Conversationally fluent*
French	**Basic**	*Basic words and phrases only*

Publications (incomplete)

PACT 2022 **Jie Zhao**, Cédric Bastoul, Yanzhi Yi, Jiahui Hu, Wang Nie, Renwei Zhang, Zhen Geng, Chong Li, Thibaut Tachon and Zhiliang Gan, *Parallelizing Neural Network Models Effectively on GPU by Implementing Reductions Atomically*, In Proceedings of the 31st International Conference on Parallel Architectures and Compilation Techniques (PACT 2022), 10-12 October, 2022, Chicago, Illinois, USA, 14 pages, to appear. (acceptance rate: 40/114=35.1%).

ICPP 2022 Xiaohan Tao, Yu Zhu, Boyang Wang, Jinlong Xu, Jianmin Pang and **Jie Zhao**, *Automatically Generating High-performance Matrix Multiplication Kernels on the Latest Sunway Processor*, In Proceedings of the 51st International Conference on Parallel Processing (ICPP 2022), 29 August - 1 September, 2022, Online Event, Bordeaux, France, 12 pages, to appear. (acceptance rate: 84/311=27.0%).

MLSys 2022 **Jie Zhao**, Xiong Gao, Ruijie Xia, Zhaochuang Zhang, Deshi Chen, Lei Chen, Renwei Zhang, Zhen Geng, Bin Cheng and Xuefeng Jin, *Apollo: Automatic Partition-based Operator Fusion through Layer by Layer Optimization*, In Proceedings of the 5th Conference on Machine Learning and Systems (MLSys 2022), 29 August - 1 September, 2022, Santa Clara, pages 1-19. (acceptance rate: 51/247=20.6%).

arXiv 2021 Jinhui Yuan, Xinqi Li, Cheng Cheng, Juncheng Liu, Ran Guo, Shenghang Cai, Chi Yao, Fei Yang, Xiaodong Yi, Chuan Wu, Haoran Zhang and **Jie Zhao**, *OneFlow: Redesign the Distributed Deep Learning Framework from Scratch*, arXiv:2110.15032v6 [cs.DC], https://doi.org/10.48550/arXiv.2110.15032, pages 1-13.

PLDI 2021 **Jie Zhao**, Bojie Li, Wang Nie, Zhen Geng, Renwei Zhang, Xiong Gao, Bin Cheng, Chen Wu, Yun Cheng, Zheng Li, Peng Di, Kun Zhang and Xuefeng Jin, *AKG: Automatic Kernel Generation for Neural Processing Units using Polyhedral Transformations*, In Proceedings of the 42nd ACM SIGPLAN Conference on Programming Language Design and Implementation (PLDI 2021), 20-25 June, 2021, Vitural, Canada, pages 1233-1248. (acceptance rate: 87/320=27.2%).

MICRO 2020 **Jie Zhao**, and Peng Di, *Optimizing the Memory Hierarchy by Compositing Automatic Transformations on Computations and Data*, In Proceedings of the 53rd IEEE/ACM International Symposium on Microarchitecture (MICRO-53), 17-21 October, 2020, Global Online Event, pages 427-441. (acceptance rate: 82/446=18.4%).

TACO 2019 **Jie Zhao** and Albert Cohen, *Flextended tiles: a flexible extension of overlapped tiles for polyhedral compilation*, ACM Transactions on Architecture and Code Optimization, 2019, 16(4): Article 47, 25 pages.

ICS 2019 HuiHui Sun, Florian Fey, **Jie Zhao**, and Sergei Gorlatch , *WCCV: Improving the Vectorization of IF-statements with Warp-coherent Conditions*, In Proceedings of the 33rd ACM International Conference on Supercomputing (*ICS'19*), 26-28 June, 2019, Phoenix, Arizona, USA, pages 319-329. (acceptance rate: 45/193=23.3%).

CC 2018 **Jie Zhao**, Micheal Kurse, and Albert Cohen, *A Polyhedral Compilation Framework for Loops with Dynamic Data-dependent Bounds*, In Proceedings of the 27th International Conference on Compiler Construction (*CC'18*), 24-25 February, 2018, Vienna, Austria, pages 14-24. (acceptance rate: 18/52=34.6%).

Book In Print

TUP Book **Jie Zhao** and Baoliang Li, *Polyhedral Compilation Theory and its Practice in Deep Learning (in Chinese)*, Tsinghua University Press, 291 pages. In print.