

田博宇 Boyu Tian

邮箱: tby20@mails.tsinghua.edu.cn 主页: <https://criust.github.io> 电话/微信: +86-15389218086

教育经历

- 直博生 清华大学交叉信息研究院 2020.9 - 2025.6 (预计)
计算机科学与技术, 导师: 高鸣宇 助理教授
- 工学学士 上海交通大学 ACM 班 2016.9 - 2020.6
计算机科学与技术, 导师: 李超 教授

研究兴趣

计算机系统与体系结构, 高效内存架构, 近存计算/存算一体, 内存池化, 分离式内存, 大模型推理加速

论文发表

- Boyu Tian, Yiwei Li, Li Jiang, Shuangyu Cai, and Mingyu Gao. NDPBridge: Enabling Cross-Bank Coordination in Near-DRAM-Bank Processing Architectures. In *ISCA*, 2024. (CCF-A)
- Boyu Tian, Qihang Chen, and Mingyu Gao. ABNDP: Co-optimizing Data Access and Load Balance in Near-Data Processing. In *ASPLOS*, 2023. (CCF-A).
- Shuangyu Cai, Boyu Tian, Huanchen Zhang, and Mingyu Gao. PimPam: Efficient Graph Pattern Matching on Real Processing-in-Memory Hardware. In *SIGMOD*, 2024. (CCF-A).
- Qihang Chen, Boyu Tian, and Mingyu Gao. FINGERS: Exploiting Fine-Grained Parallelism in Graph Mining Accelerators. In *ASPLOS*, 2022. (CCF-A).
- Bohan Zhao, Xiang Li, Boyu Tian, Zhiyu Mei, and Wenfei Wu. DHS: Adaptive Memory Layout Organization of Sketch Slots for Fast and Accurate Data Stream Processing. In *KDD*, 2021. (CCF-A)

科研经历

- 清华大学交叉信息研究院 IDEAL Lab 2020.9 - 至今
博士生, 导师: 高鸣宇 助理教授 中国, 北京
- 大规模近存计算系统的性能优化, 包括数据通信支持和负载均衡等。
 - 基于近存计算硬件和 GPU 的异构大模型推理系统设计和优化。
 - 基于 CXL 的数据中心上可扩展、可支持异构设备和异构内存的大规模内存池。
- 上海交通大学电子信息与电气工程学院 SAIL Lab 2018.7 - 2020.6
研究实习生, 导师: 李超教授 中国, 上海
- 可根据 QoS 要求进行动态调度的近似图计算系统。
 - 数据中心针对微服务应用的动态资源调度。
- 杜克大学 CEI Lab 2019.7 - 2019.9
研究实习生, 导师: 陈怡然教授 美国, 北卡罗来纳
- 基于非易失性内存存内计算的图计算加速器。

实习经历

智谱 AI	2024.3 - 至今
研究实习生，导师：冯冠宇博士	中国，北京
<ul style="list-style-type: none">大模型推理加速框架开发。针对大模型推理的新硬件设计和优化探索。商业化近存计算硬件上的大模型推理技术探索。	
阿里巴巴达摩院	2023.6 - 2024.1
研究实习生，导师：牛迪民博士	中国，北京
<ul style="list-style-type: none">首个可商业化的基于 CXL 真实硬件的数据中心内存池设计和开发。可扩展、可支持异构硬件的 CXL 内存池设计和仿真。	
华为海思图灵架构与设计部	2019.10 - 2019.12
研究实习生，导师：廖恒博士、李琳博士	中国，上海
<ul style="list-style-type: none">基于稀疏的图像输入的 3D 视频合成方案。	

获奖情况

清华大学综合优秀奖学金	2021, 2022, 2023
ASPLOS 2023 Student Travel Award	2023
唐立新奖学金	2018-2020
上海交通大学致远杰出领袖奖学金	2017
上海交通大学致远荣誉奖学金	2016-2019

教学经历

计算机系统结构 助教	清华大学 2021 春, 2022 春
C++ 程序设计 助教	上海交通大学 2017 秋

技术能力

编程语言	C, C++, Python, Verilog, Java, Rust, Go
硬件模拟框架	ZSim, Intel Pin, CACTI, Ripes
大模型推理框架	Text Generation Inference