

翁楚良

Email: chuliang.weng@gmail.com

电话: +86 189 1653 8856

<http://chuliangweng.github.io>

研究方向

并行与分布式系统：智能数据处理系统、虚拟化技术与云计算、内存计算系统

学习与工作经历

2016.05 - 现在，教授/博士生导师，华东师范大学

2015.01 - 2018.10，科技部 **863** 项目负责人，华为技术有限公司

2013.05 - 2016.04，首席/主任研究员，华为中央研究院

2011.11 - 2012.11，Visiting Research Scientist，哥伦比亚大学

2004.09 - 2013.05，讲师/副教授，上海交通大学

2001.03 - 2004.06，博士，上海交通大学

1994.09 - 2001.02，本科/硕士，西南交通大学

项目节选

围绕并行与分布式系统，针对人工智能数据处理系统、内存计算与新型存储系统、虚拟化技术与云计算等开展研发，取得系列成果。

1. 数据处理平台与技术 (2016 -)

1) 结合自身工作经历，从工业界回到高校后开展数据处理全栈研究，贯穿数据处理应用层、操作系统层、体系结构层。针对大数据系统、深度学习系统中的数据处理，从计算加速 (GPU/SIMD)、存储加速 (NVM/In-Memory) 和分布式网络 (RDMA/DPDK)、及算法模型等上开展研究工作，逐步构建面向大数据和机器学习的数据处理平台关键技术。

2) 在计算方面，结合处理器 SIMD 硬件功能，采用软硬件相结合的方法实现数据处理的加速，基于 Intel CPU AVX512 指令集实现数据处理加速的阶段性工作发表于 VLDB2020^[5]，进一步的工作包括从协程程序库 (coroutine) 和处理器硬件两个层面逐步落地。基于 NVIDIA GPU 的硬件缺页等功能，通过 GPU 传输数据块预选、GPU 计算和传输叠加等实现在 GPU 上数据处理加速，支持粗粒度和细粒度的 Pipeline，成果发表在顶级期刊 IEEE Transactions on Computers^[7]。通过技术面和数据面的分离技术，解决了目前数据平台中无法有效利用 CPU+GPU 的混合算力问题，成果发表在并行计算顶级会议 PPOPP2021^[2]。

3) 在存储方面，基于 NVMe 存储特性，针对大数据系统以表对象为主、深度学习系统以 DataFrame 为主的特征，提出并实现了原生的维度表存储，单机版工作发表在顶级会议 ICDE2020^[6]，结合 RDMA 的分布式环境下负载均衡工作发表于 ICDE2021^[3]。原生表存储是与文件存储、块存储、对象存储等基本存储不同的一种全新存储数据结构和系统。进一步的研究正在阿里巴巴创新研究计划 (AIR) 项目资助下进行。

4) 在系统平台方面,基于在计算和存储方面的工作成果,实现自主的数据处理系统,支持大数据实时注入、实时分析和智能推断,相关工作开源 <https://github.com/daseECNU/Ginkgo>。结合证券数据分析应用,分布式环境查询加速和优化工作发表在顶级期刊 IEEE TKDE 等上^[8],实时事务型数据注入研究工作被 IEEE TKDE 录用^[4],实时智能推断工作已完成阶段性的研究并投稿,目前在开展面向大数据的分布式机器学习平台技术研究。

2. 新型数据存储与硬件加速 (2013 –)

1) 应邀加入华为中央研究院,担任主任研究员/首席研究员 (2013–2016),针对数据中心热数据高性能存储的需求,研究新型非易失内存 (NVM) 技术。作为项目负责人,完成该项研究的整体规划和研究布局 (包括新型存储控制器、存储软件、整机架构等),组建并带领国际化团队 (包括美国硅谷) 研制智能内存存储系统原型,提出的单式存储技术成果写入华为公司年报。技术落地华为 IT 产品线,大幅提升 SAP HANA 应用性能。

2) 作为华为公司国家项目负责人 (2015–2018),负责国家科技部 863 计划项目“面向大数据处理的混合内存硬件平台架构设计与开发”,组织和带领来自华为公司、上海交通大学和国防科技大学等单位的项目成员,完成新型非易失内存 (NVM) 服务器样机及大数据智能加速硬件系统的研制,在国际标准组织 JEDEC 通过 6 项 NVDIMM-P 标准提案。2018 年 10 月科技部专家组在华为对该 863 计划项目进行了验收,认定混合内存关键技术验证平台突破了当前国外技术壁垒。目前结合混合内存 (DRAM + Optane DIMM) 在开展大数据系统和深度学习系统中冷热数据存储的研究工作。

3. 云计算与虚拟化技术 (2009 –)

1) 在上海交通大学计算机系工作期间于 2011 年到美国哥伦比亚大学计算机系软件系统实验室与 Junfeng Yang 教授开展为期一年的访问研究,围绕虚拟化系统中多线程程序性能优化开展研究。相关成果发表在计算机系统顶级期刊 IEEE Transactions on Computers^{[9][10][11]},实测性能优于 VMware 产品的同类型高级功能特性。

1) 在华为中央研究院工作期间,落地云计算与虚拟化技术。其中,提出的虚拟 CPU 协同调度方法获得华为 IT 产品线的谈判专利奖 (华为在云计算领域年度唯一);带领华为国内外团队针对云平台中虚拟机提出了防止内存数据泄露技术。目前在开展大型虚拟化系统性能隔离和云主机系统安全的研究,研究工作获国家自然科学基金项目资助,相关成果被信息安全顶级期刊 IEEE TDSC 录用^[1]。

4. 科研项目节选

1) 国家自然科学基金面上项目“大型虚拟化系统性能隔离强化机制研究”,项目负责人,项目批准号: 61772204,时间: 2018.01-2021.12,直接经费: 65 万元)

2) 阿里巴巴创新研究计划 (AIR) 项目“云原生数据库备份的存储系统研究”,项目负责人,经费: 50 万元,时间: 2021.01-2022.01

3) 科技部 863 计划项目“面向大数据处理的混合内存硬件平台架构设计与开发”,项目负责人,项目批准号: No. 2015AA015302,经费: 2927 万元,时间: 2015.01-2017.12

4) 国家自然科学基金重大研究计划培育项目“重大工程动力灾变模拟系统集成的关键技术究”,项目负责人,经费: 50 万元,项目批准号: 90715030,时间: 2008.01-2010.12

5) 国家自然科学基金面上项目“虚拟化系统自适应调度关键技术究”,项目负责人,经费: 30 万元,项目批准号: 60970008,时间: 2010.01-2012.12

研究论文

1. 近年 CCF-A 类论文节选 (第一作者或通讯作者, 均为国内团队独立完成)

- [1] Beilei Zheng, Jianan Gu, Jialun Wang, and **Chuliang Weng***. CBA-Detector: A Self-Feedback Detector against Cache-Based Attacks. *IEEE Transactions on Dependable and Secure Computing (TDSC)*, 2021, Regular paper, Accepted
- [2] Zhifang Li, Mingcong Han, Shangwei Wu, and **Chuliang Weng***. ShadowVM: Accelerating Data Plane for Data Analytics with Bare Metal CPUs and GPUs. *The 26th ACM SIGPLAN Annual Symposium on Principles and Practice of Parallel Programming (PPoPP)*, Korea, pp. 147-160, 2021
- [3] Zeyu He, Zhifang Li, Xiaoshuang Peng, and **Chuliang Weng***. DS²: Handling Data Skew in Spark SQL Using Data Stealing over High-Speed Networks. *The 37th IEEE International Conference on Data Engineering (ICDE)*, 2021, Short paper, Accepted
- [4] Zhifang Li, Beicheng Peng, Qiuli Huang, and **Chuliang Weng***. Karst: Transactional Data Ingestion without Blocking on a Scalable Architecture. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 2020, Regular paper, Accepted
- [5] Zhuhe Fang, Beilei Zheng, and **Chuliang Weng***. Interleaved MultiVectorizing. *The 46th International Conference on Very Large Data Bases (VLDB)*, Japan, 2020, Proceedings of the VLDB Endowment, 2019, 13(3): 226-238
- [6] Jiajia Chu, Yunshan Tu, Yao Zhang, and **Chuliang Weng***. LATTE: A Native Table Engine on NVMe Storage. *The 36th IEEE International Conference on Data Engineering (ICDE)*, Dallas, Texas, USA, pp. 1225-1236, 2020
- [7] Zhifang Li, Beicheng Peng, and **Chuliang Weng***. XeFlow: Streamlining Inter-processor Pipeline Execution for the Discrete CPU-GPU Platform. *IEEE Transactions on Computers (TC)*, 2020, 69(6):819-831
- [8] Zhuhe Fang, **Chuliang Weng***, Li Wang, Huiqi Hu, and Aoying Zhou. Scheduling Resources to Multiple Pipelines of One Query in a Main Memory Database Cluster. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, 2020, 32(3): 533-546
- [9] **Chuliang Weng***, Qian Liu, Kenli Li, and Deqing Zou. CloudMon: Monitoring Virtual Machines in Clouds. *IEEE Transactions on Computers (TC)*, 2016, 65(12): 3787-3794
- [10] **Chuliang Weng***, Jianfeng Zhan, and Yuan Luo. TSAC: Enforcing Isolation of Virtual Machines in Clouds. *IEEE Transactions on Computers (TC)*, 2015, 64(5): 1470-1482
- [11] **Chuliang Weng***, Minyi Guo, Yuan Luo, and Minglu Li. Hybrid CPU Management for Adapting to the Diversity of Virtual Machines. *IEEE Transactions on Computers (TC)*, 2013, 62(7): 1332-1344

2. 国内外合作论文节选

- [12] Jianguo Chen, Kenli Li, Zhuo Tang, Kashif Bilal, Shui Yu, **Chuliang Weng**, and Keqin Li. A Parallel Random Forest Algorithm for Big Data in a Spark Cloud Computing Environment. IEEE Transactions on Parallel and Distributed Systems (*TPDS*), 2017, 28(4): 919-933
- [13] Yuhao Wang, Hao Yu, Leibin Ni, Mei Yan, Guangbin Huang, **Chuliang Weng**, Wei Yang, and Junfeng Zhao. An Energy-efficient Nonvolatile In-memory Computing Architecture for Extreme Learning Machine by Domain-wall Nanowire Devices, IEEE Transactions on Nanotechnology (*TNANO*), 2015, 14(6):998-1012
- [14] Yuhao Wang, Hantao Huang, Leibin Ni, Hao Yu, Mei Yan, **Chuliang Weng**, Wei Yang, and Junfeng Zhao. An Energy-efficient Non-volatile In-Memory Accelerator for Sparse-representation based Face Recognition. The 18th Design, Automation and Test in Europe (*DATE*), 2015
- [15] Gang Lu, Jianfeng Zhan, Haining Wang, Lin Yuan, Yunwei Gao, **Chuliang Weng**, and Yong Qi. PowerTracer: Tracing Requests in Multi-tier Services to Reduce Energy Inefficiency. IEEE Transactions on Computers (*TC*), 2015, 64(5): 1389-1401
- [16] Jianfeng Zhan, Lei Wang, Xiaona Li, Weisong Shi, **Chuliang Weng**, Wen Yao Zhang, and Xiutao Zang. Cost-aware Cooperative Resource Provisioning for Heterogeneous Workloads in Data Centers. IEEE Transactions on Computers (*TC*), 2013, 62(11): 2155-2168

获奖节选

- [1] 2020, “面向内存计算的软件定义关键技术及应用”, 上海市技术发明奖一等奖 (排名 5/14)
- [2] 2016, 华为公司 IT 产品线谈判专利奖 (第一发明人, 华为公司云计算领域年度唯一)
- [3] 2014, 带领的“下一代计算/存储系统研究团队”, 获得华为公司 2012 实验室总裁奖
- [4] 2010, 第七届中国通信学会学术年会优秀论文奖
- [5] 2008, 上海交通大学晨星青年学者奖励计划优秀青年教师奖一等奖