

# 陈游旻

<http://storage.cs.tsinghua.edu.cn/~cym>

北京市海淀区清华大学东主楼 8 区 201 室, 100084 | 13121567386 | chenym16@mails.tsinghua.edu.cn

## 个人介绍

我的主要研究方向为分布式系统和存储系统。博士期间, 我的主要研究工作包括: 1) 基于新型硬件 (如 RDMA、持久性内存、可编程网络等) 构建高性能 (分布式) 文件系统、键值存储系统、事务系统等; 2) 解决数据中心尾延迟问题, 成功在持久性数据索引、内存事务系统等场景下将尾延迟降低一个数量级; 3) 新型持久性内存在数据中心的硬件安全问题等。相关论文发表至 FAST、ASPLOS、VLDB、ICDE、Eurosys、USENIX ATC、DAC 等国际顶级会议及 IEEE TC、ACM TOS 等国际著名期刊。其中, CCF 推荐列表或清华计算机系推荐列表 (TH-CPL) 中 A 类论文共发表 14 篇, 以第一作者 (或学生一作) 身份发表论文共 8 篇。

## 教育背景

**留学基金委博士生联合培养项目 | 2019 年 11 月至 2020 年 11 月 | 美国威斯康星大学麦迪逊分校**

- 院系: 计算机系, 高级系统实验室 (The ADvanced Systems Laboratory, ADSL)
- 导师: Remzi Arpaci-Dusseau, Andrea Arpaci-Dusseau, Xiangyao Yu

**博士 | 2016 年 8 月至 2021 年 6 月 | 清华大学**

- 院系: 计算机系
- 导师: 舒继武教授

**本科 | 2012 年 8 月至 2016 年 6 月 | 北京航空航天大学**

- 院系: 沈元荣誉学院 (从全校范围选拔各系前 3 名, 本人专业为计算机科学与技术)

## 发表论文

### 投稿中

1. **Youmin Chen**, Xiangyao Yu, Andrea C. Arpaci-Dusseau, Remzi H. Arpaci-Dusseau, and Jiwu Shu. “High Throughput and Low Tail Latency Transactions with Plor” (In submission to **OSDI’21, CCF-A**)

### 一作或学生一作身份发表论文

2. **Youmin Chen**, Youyou Lu, Bohong Zhu, Andrea C. Arpaci-Dusseau, Remzi H. Arpaci-Dusseau, and Jiwu Shu. “KucoFS: a Kernel/User-level Collaborative File System for Non-Volatile Memories” (**FAST’21, CCF-A**), 2021.
3. **Youmin Chen**, Youyou Lu, Kedong Fang, Qing Wang, and Jiwu Shu. “ $\mu$ Tree: a Persistent B+-Tree with Low Tail Latency”, 46th International Conference on Very Large Data Bases (**VLDB’20, CCF-A**), 2020.
4. **Youmin Chen**, Youyou Lu, Fan Yang, Qing Wang, Yang Wang, and Jiwu Shu. “FlatStore: an Efficient Log-Structured Storage Engine for Persistent Memory Key-value Stores”, Proceedings of the 25th International

- Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS'20, CCF-A), 2020.
5. **Youmin Chen**, Youyou Lu, and Jiwu Shu. “Scalable RDMA RPC on Reliable Connection with Efficient Resource Sharing”, European Conference on Computer Systems (Eurosys'19, TH-CPL-A, CCF-B), 2019.
  6. Jiwu Shu, **Youmin Chen**, Qing Wang, Bohong Zhu, Junru Li, and Youyou Lu. “TH-DPMS: Design and Implementation of an RDMA-enabled Distributed Persistent Memory Storage System”, ACM Transactions on Storage (TOS, CCF-A, 已录用, 学生一作, 2020).
  7. Jiwu Shu, **Youmin Chen**, Qingda Hu, and Youyou Lu. “The Development of System Software on Non-volatile Main Memory”. SCIENTIA SINICA Informationis, 2020. (SSI, TH-CPL-A, 已录用, 学生一作)
  8. **Youmin Chen**, Youyou Lu, Pei Chen, and Jiwu Shu. “Efficient and Consistent NVMM Cache for SSD-based File System”, IEEE Transactions on Computers (TC, CCF-A), (2018).
  9. **Youmin Chen**, Jiwu Shu, Jiaxin Ou, and Youyou Lu. “HiNFS: A persistent memory file system with both buffering and direct-access”, ACM Transactions on Storage (TOS, CCF-A) 14.1 (2018): 4.
  10. **Youmin Chen**, Youyou Lu, Shengmei Luo, and Jiwu Shu. “Survey on RDMA-based Distributed Storage Systems”, Journal of Computer Research and Development, 2019, 56(2): 227-239. (TH-CPL-B, 中文).
  11. **Youmin Chen**, Bohong Zhu, Yinjun Han, Yaofeng Tu, and Jiwu Shu. “A Hybrid Approach for Managing Data Pages in Persistent Memory File Systems”. Journal of Computer Research and Development (TH-CPL-B, 中文), 2020.
  12. 陈游旻, 李飞, 舒继武, “大数据环境下的存储系统构建: 挑战、方法和趋势”, 《大数据》, 5 (4) : 27-40, 2019 (中文)。

## 非一作发表论文

13. Fan Yang, **Youmin Chen**, Youyou Lu, Qing Wang, Jiwu Shu. Aria: Tolerating Skewed Workloads in Secure In-memory Key-value Stores. In Proceedings of the 37th IEEE International Conference on Data Engineering (ICDE' 21, CCF-A), April 19-22, 2021 (online).
14. Qing Wang, Youyou Lu, Junru Li, **Youmin Chen**, Erci Xu, and Jiwu Shu. “Concordia: Distributed Shared Memory with In-Network Cache Coherence” (FAST'21, CCF-A), 2021
15. Fan Yang, Youyou Lu, **Youmin Chen**, and Jiwu Shu. “No Compromises: Secure NVM with Crash Consistency, Write-Efficiency and High-Performance”, Design Automation Conference (DAC'19, CCF-A), 2019.
16. Youyou Lu, Jiwu Shu, **Youmin Chen**, and Tao Li. “Octopus: an RDMA-enabled Distributed Persistent Memory File System”, USENIX Annual Technical Conference (USENIX ATC'17, CCF-A, 学生一作), 2017.
17. Fan Yang, **Youmin Chen**, Haiyu Mao, Youyou Lu, and Jiwu Shu. ShieldNVM: An Efficient and Fast Recoverable System for Secure Non-Volatile Memory. ACM Transactions on Storage (TOS, CCF-A), 2020.
18. Bohong Zhu, **Youmin Chen**, Qing Wang, Youyou Lu, Jiwu Shu. Octopus+: an RDMA-enabled Distributed Persistent Memory File System. ACM Transactions on Storage (ToS, CCF-A), 已录用, 2021.
19. Jiangkun Hu, **Youmin Chen**, Youyou Lu, Xubin He, Jiwu Shu. Understanding and analysis of B+-trees on NVM towards consistency and efficiency. CCF Transactions on High Performance Computing (CCF-HPC), vol. 2, pp, 36-49, 2020.
20. 陈波, 陆游游, 蔡涛, 陈游旻, 屠要峰, 舒继武, 一种分布式持久性内存文件系统的一致性机制, 《计算机研究与发展》, 57 (3): 660-667, 2020. (TH-CPL-B)
21. 陈娟, 胡庆达, 陈游旻, 陆游游, 舒继武, 杨晓辉, 一种基于微日志的持久性事务内存系统, 《计算机研究与发展》, 55 (9): 2029-2037, 2018. (TH-CPL-B)

## WORKSHOP 及摘要论文

22. Youyou Lu, Jiwu Shu, **Youmin Chen**, Tao Li. “Octopus: an RDMA-enabled Distributed Persistent Memory File System”, Non-Volatile Memory Workshop (NVMW'19). 2019.

23. **Youmin Chen**, Jiwu Shu, Youyou Lu “Scalable and Reliable RDMA”, Proceedings of the ACM Symposium on Cloud Computing (SoCC’18). ACM, 2018.

## 申请及授权的发明专利

### 已授权专利

1. 舒继武, **陈游旻**, 李飞, 陆游游, 分布式持久性内存存储系统的构建方法, 中国发明专利授权号: CN110221779B, 授权公告日: 2020.06.19。
2. 陆游游, 舒继武, **陈游旻**, 一种基于 RDMA 的分布式内存文件系统, 中国发明专利授权号: CN108268208B, 授权公告日: 2020.01.17。
3. 陆游游, 舒继武, **陈游旻**, 一种基于 RDMA 的高并发数据传输方法, 中国发明专利授权号: CN106657365B, 授权公告日: 2019.12.17。

### 公开未授权专利

4. Jiwu Shu, **Youmin Chen**, Bohong Zhu, Youyou Lu, DATA STORAGE ACCESS METHOD, DEVICE AND APPARATUS FOR PERSISTENT MEMORY. (美国专利, 申请号: 16553276)
5. Jiwu Shu, **Youmin Chen**, Bohong Zhu, Youyou Lu, PERSISTENT MEMORY STORAGE ENGINE DEVICE BASED ON LOG STRUCTURE AND CONTROL METHOD. (美国专利, 申请号: 16553253)
6. 陆游游, 舒继武, **陈游旻**, 陈佩, 徐君, 林芃, 一种基于远程直接内存访问 RDMA 的内存通信方法及装置, 中国发明专利申请公布号: CN111858418A, 申请公布日: 2020.10.30。
7. 舒继武, 汪庆, 陆游游, **陈游旻**, 一种分布式持久性内存事务系统的日志管理方法, 中国发明专利申请公布号: CN111858418A, 申请公布日: 2020.10.30。
8. 舒继武, 唐小岚, 陆游游, **陈游旻**, 杨洪章, 张晗, 一种基于 RDMA 的数据传输方法和分布式共享内存系统, 中国发明专利申请公布号: CN111277616A, 申请公布日: 2020.06.12。
9. 舒继武, **陈游旻**, 朱博弘, 陆游游, 一种持久性内存对象存储系统, 中国发明专利申请公布号: CN111240588A, 申请公布日: 2020.06.05。
10. 舒继武, **陈游旻**, 汪庆, 陈佩, 陆游游, 一种分布式持久性内存的全局地址空间管理方法, 中国发明专利申请公布号: CN111241011A, 申请公布日: 2020.06.05。
11. 舒继武, **陈游旻**, 朱博弘, 陆游游, 持久性内存的数据存储访问方法、设备及装置, 中国发明专利申请公布号: CN110377436A, 申请公布日: 2019.10.25。
12. 舒继武, **陈游旻**, 朱博弘, 陆游游, 基于日志结构的持久性内存存储引擎装置及控制方法, 中国发明专利申请公布号: CN110377531A, 申请公布日: 2019.10.25。
13. 舒继武, **陈游旻**, 陆游游, 崔文林, 读写请求处理方法、装置、电子设备以及存储介质, 中国发明专利申请公布号: CN109885393A, 申请公布日: 2019.06.14。

## 获奖情况

- 中国计算机协会科学技术奖技术发明一等奖 (个人排名第 2) | 中国计算机学会 | 2020
- 奥林帕斯奖及百万悬红 (个人排名第 3) | 华为 | 2020
- 国家奖学金 | 教育部 | 2020
- 一等奖学金 | 清华大学 | 2018/2019

- 瑞安奖学金 | 北京航空航天大学 | 2015
- 学习优秀奖二等奖 | 北京航空航天大学 | 2015
- 科技竞赛一等奖 | 北京航空航天大学 | 2015
- 全国大学生数学建模竞赛北京市一等奖 | 北京航空航天大学 | 2015
- 新生入学奖学金二等奖 | 北京航空航天大学 | 2012

## 学术兼职

- 担任国际著名期刊 ACM Transactions on Storage (TOS)、Journal of Systems and Architecture、IEEE Transactions on Computers (TC) 审稿人
- 受邀加入新创期刊 Journal of Systems Research (JSys) 学生编委会

## 项目研究经历

### 作为研究骨干参与的科研项目

1. 科技部重点研发计划：TB 级持久性内存存储技术与系统（2018YFB1003301），2018.8 - 2021.7。
2. 国家自然科学基金重点项目：新型分布式存储系统的高可靠性关键技术研究（61832011），2019.1 - 2023.12。
3. 华为技术有限公司项目：持久性内存与智能存储系统（YBN2019125112），2020.1~ 2022.12。
4. 中兴通讯股份有限公司课题：基于 NVM 的分布式混合存储系统，2018.11 - 2020.11

### 研究项目简介

#### OCTOPUS – 基于 RDMA 的分布式持久性内存文件系统

- 完成时间：2015 年-2016 年
- 我和陆游游副教授共同设计系统，并由我完成系统实现及实验验证。
- 持久性内存及 RDMA 网络提供了极高的硬件性能，但是现有的分布式文件系统并不能充分发挥其硬件性能。Octopus 抽象出了统一的共享内存池并重新设计了文件系统内部机制，从而提供了接近硬件的文件访问性能。Octopus 是国际上**最早**研制出的基于持久性内存及 RDMA 的分布式文件系统。相关论文发表至国际顶级会议 USENIX ATC，目前已拥有近 70 次引用。

#### SCALERPC – 优化 RDMA 扩展性

- 完成时间：2016 年-2017 年
- 本人独立完成系统设计、实现及实验验证。
- RDMA 在海量连接下具有扩展性差的问题，这主要是因为 RDMA 网卡及 CPU 的缓存空间受限，海量连接数导致缓存空间竞争，从而使得总体吞吐下降。ScaleRPC 是一个基于 RMDA 的远程过程调用系统，它通过硬件资源复用及高效调度完美解决了 RDMA 扩展性问题。

#### FLATSTORE – 面向新型介质的高效键值存储系统

- 完成时间：2017 年-2018 年
- 本人独立完成系统设计，杨帆和汪庆协助系统实现及实验验证。

- 英特尔的持久性内存 Optane DCPMM 已经于 2019 年 4 月正式发布。然而，Optane 虽具备字节寻址能力，但是其内部的更新粒度高达 256 字节，这远大于键值存储系统中典型的更新粒度。硬件更新粒度与应用更新粒度不匹配的问题使得 Optane 的硬件利用率极低。FlatStore 针对这一问题提出了面向日志结构的跨核批量持久化机制，使得合并多次更新以降低持久性内存写入次数，同时不影响延迟。

### **KUCOFS – 基于持久性内存的高性能文件系统**

- 完成时间：2018 年-2019 年
- 本人独立完成系统设计，朱博弘协助系统实现和实验验证。
- 近年来，不少研究者提出在用户态构建文件系统以直接访问持久性内存，从而降低内核态的软件开销。然而，用户态文件系统往往需要引入第三方进程更新核心元数据，在高并发场景下，第三方进程很容易成为性能瓶颈，导致扩展性问题。KucoFS 是一个用户态和内核态协同工作的文件系统，其中，元数据通过内核态线程更新，而数据读写则直接在用户态完成。同时，为避免内核线程成为性能瓶颈，引入协同索引、多级并发控制等概念，将耗时的操作从内核态卸载到用户态，进一步提升扩展性。

### **CHRONUS – 事务系统的尾延迟优化**

- 完成时间：2019 年-2020 年
- 本人独立完成系统设计、系统实现和实验验证。
- 尾延迟问题已经成为数据中心越来越关注的问题。近年，不少的研究已经从操作系统不同层次尝试解决尾延迟问题，然而，请求冲突造成的尾延迟问题却极少得到关注。为回答这一问题，本项目提出了“悲观锁+乐观读”的事务并发控制概念，成功将尾延迟降低一个数量级。

## **受邀报告**

### **OCTOPUS: AN RDMA-ENABLED DISTRIBUTED PERSISTENT MEMORY FILE SYSTEM**

- 2017, Santa Clara, CA, USA
- 2018, UC San Diego, CA, USA

### **SCALERPC: SCALABLE RDMA RPC ON RELIABLE CONNECTION WITH EFFICIENT RESOURCE SHARING**

- 2019, Dresden, Germany

### **FLATSTORE: AN EFFICIENT LOG-STRUCTURED STORAGE ENGINE FOR PERSISTENT MEMORY KEY-VALUE STORES**

- 2020, Lausanne, Switzerland, online

### **UTREE, A PERSISTENT B+-TREE WITH LOW TAIL LATENCY**

- 2020, Tokyo, Japan, online