Huaicheng LI

Contact

Phone Number: (540) 231-4482 — Email: huaicheng@cs.vt.edu
Gilbert Place 4109, Blacksburg, VA 24060 — Website: https://people.cs.vt.edu/~huaicheng

Research Interests

Areas: Operating Systems, Storage and Memory Systems, Systems Support for Emerging Hardware/Applications
Focus: Design and build novel computing systems for emerging storage/memory hardware to achieve

1. Performance: Software/Hardware co-design for low and predictable end-to-end latencies and high throughput
2. Efficiency: Offloaded and disaggregated system architecture designs for improved resource and cost efficiency
3. Programmability: Systems support for emerging I/O and acceleration technologies to ease development efforts

Academic Positions

Virginia Tech — Blacksburg, VA
Assistant Professor, Department of Computer Science — 2022–Present

Carnegie Mellon University — Pittsburgh, PA
Postdoctoral Researcher, Parallel Data Lab (PDL) — 2020–2022
Supervisor: Gregory R. Ganger

Education

University of Chicago — Chicago, IL
Ph.D. in Computer Science (*M.S. conferred in 2018*) — 2020
Advisor: Haryadi S. Gunawi
Thesis: Evolving Storage Stack for Predictability and Efficiency [D1]

Wuhan University — Wuhan, China
B.S. in Computer Science and Technology — 2013

Honors & Awards

NSF CAREER Award	2024
IEEE Micro Top Picks 2024 Honorable Mention: Pond [C1]	2024
ASPLOS'23 Distinguished Paper Award: Pond [C1]	2023
SYSTOR'22 Best Paper Award: Fantastic SSD Internals [C3]	2022
SYSTOR'21 Distinguished Reviewer Award	2021
Nomination for the SIGOPS Dennis M. Ritchie Doctoral Dissertation Award (*1 per department*)	2020
University Unrestricted (UU) Fellowship, University of Chicago	2019
FAST'18 Best Paper Nominee: Fail-Slow at Scale [C7]	2018
FAST'17 Best Paper Nominee: Tiny-Tail Flash [C9]	2017

Conference Publications

Bibliometrics on Google Scholar and DBLP

ASPLOS'23 [C1] **Huaicheng Li**, Daniel S. Berger, Stanko Novakovic, Lisa Hsu, Daniel Ernst, Pantea Zardoshti, Monish Shah, Ishwar Agarwal, Mark D. Hill, Marcus Fontoura, Ricardo Bianchini. **Pond: CXL-Based Memory Pooling Systems for Cloud Platforms**. In the Proceedings of the 28th ACM International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS), 2023.
Distinguished Paper Award
IEEE Micro Top Picks 2024 Honorable Mention

ASPLOS'23 [C2] Thomas Kim, Jekyeom Jeon, Nikhil Arora, **Huaicheng Li**, Michael Kaminsky, David G. Andersen, Gregory R. Ganger, George Amvrosiadis, Matias Bjørling. **RAIZN: Redundant Array of Independent Zoned Namespaces**. In the Proceedings of the 28th ACM International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS), 2023.

SYSTOR'22 [C3] Nanqinqin Li, Mingzhe Hao, **Huaicheng Li**, Xing Lin, tim Emami, Haryadi S. Gunawi. **Fantastic SSD Internals and How to Learn and Use Them.** In the Proceedings of the 15th ACM International Systems and Storage Conference (SYSTOR), 2022.
Best Paper Award

SOSP'21 [C4] **Huaicheng Li**, Martin L. Putra, Ronald Shi, Xing Lin, Gregory R. Ganger, Haryadi S. Gunawi. **IODA: A Host/Device Co-Design for Strong Predictability Contract on Modern Flash Storage.** In the Proceedings of the 28th Symposium on Operating Systems Principles (SOSP), 2021.

ASPLOS'20 [C5] **Huaicheng Li**, Mingzhe Hao, Stanko Novakovic, Vaibhav Gogte, Sriram Govindan, Dan R. K. Ports, Irene Zhang, Ricardo Bianchini, Haryadi S. Gunawi, Anirudh Badam. **LeapIO: Efficient and Portable Virtual NVMe Storage on ARM SoCs.** In the Proceedings of the 25th ACM International Conference on Architectural Support for Programming Languages and Operating Systems (ASPLOS), 2020.

FAST'18 [C6] **Huaicheng Li**, Mingzhe Hao, Michael Hao Tong, Swaminatahan Sundararaman, Matias Bjørling, Haryadi S. Gunawi. **The CASE of FEMU: Cheap, Accurate, Scalable and Extensible Flash Emulator.** In the Proceedings of the 16th USENIX Conference on File and Storage Technologies (FAST), 2018.

FAST'18 [C7] Haryadi S. Gunawi, Riza Suminto, Russell Sears, Casey Golliher, Swaminatahan Sundararaman, Xing Lin, tim Emami, Weiguang Sheng, Nematollah Bidokhti, Caitie McCaffrey, Gary Grider, Parks M. Fields, Kevin Harms, Robert B. Ross, Andree Jacobson, Robert Riccio, Kirk Webb, Peter Alvaro, H. Birali Runesh, Mingzhe Hao, **Huaicheng Li**. **Fail-Slow at Scale: Evidence of Hardware Performance Faults in Large Production Systems.** In the Proceedings of the 16th USENIX Conference on File and Storage Technologies (FAST), 2018.
Best Paper Nominee

SOSP'17 [C8] Mingzhe Hao, **Huaicheng Li**, Michael Hao Tong, Chrisma Pakha, Riza Suminto, Cesar A. Stuardo, Andrew A. Chien, Haryadi S. Gunawi. **MittOS: Supporting Millisecond Tail Tolerance with Fast Rejecting SLO-Aware OS Interface.** In the Proceedings of the 26th Symposium on Operating Systems Principles (SOSP), 2017.

FAST'17 [C9] Shiqin Yan, **Huaicheng Li**, Mingzhe Hao, Michael Hao Tong, Swaminatahan Sundararaman, Andrew A. Chien, Haryadi S. Gunawi. **Tiny-Tail Flash: Near-Perfect Elimination of Garbage Collection Tail Latencies in NAND SSDs.** In the Proceedings of the 15th USENIX Conference on File and Storage Technologies (FAST), 2017.
Best Paper Nominee

Journal Publications

IEEE Micro'23 [J1] Daniel S. Berger, Daniel Ernst, **Huaicheng Li**, Pantea Zardoshti, Monish Shah, Samir Rajadnya, Scott Lee, Lisa Hsu, Ishwar Agarwal, Mark D. Hill, Ricardo Bianchini. **Design Tradeoffs in CXL-Based Memory Pools for Cloud Platforms.** IEEE Micro Special Issue on Emerging System Interconnects, 2023.

TOS'23 [J2] **Huaicheng Li**, Martin L. Putra, Ronald Shi, Fadhil I. Kurnia, Xing Lin, Jaeyoung Do, Achmad I. Kistijantoro, Gregory R. Ganger, Haryadi S. Gunawi. **Extending and Programming the NVMe I/O Determinism Interface for Flash Arrays.** ACM Transactions on Storage (TOS), Volume 19, Issue 1, February 2023. [Extended version of C4]

TOS'18 [J3] Haryadi S. Gunawi, Riza Suminto, Russell Sears, Casey Golliher, Swaminatahan Sundararaman, Xing Lin, tim Emami, Weiguang Sheng, Nematollah Bidokhti, Caitie McCaffrey, Gary Grider, Parks M. Fields, Kevin Harms, Robert B. Ross, Andree Jacobson, Robert Riccio, Kirk Webb, Peter Alvaro, H. Birali Runesh, Mingzhe Hao, **Huaicheng Li**. **Fail-Slow at Scale: Evidence of Hardware Performance Faults in Large Production Systems.** ACM Transactions on Storage (TOS), Volume 14, Issue 3, November 2018. [Extended version of C7]
Fast-tracked

TOS'17 [J4] Shiqin Yan, **Huaicheng Li**, Mingzhe Hao, Michael Hao Tong, Swaminatahan Sundararaman, Andrew A. Chien, Haryadi S. Gunawi. **Tiny-Tail Flash: Near-Perfect Elimination of Garbage Collection Tail Latencies in NAND SSDs.** ACM Transactions on Storage (TOS), Volume 13, Issue 3, October 2017. [Extended version of C9]
Fast-tracked

Workshop Publications

HotStorage'23 [W1] Dongjoo Seo, Ping-Xiang Chen, **Huaicheng Li**, Matias Bjørling, Nikil Dutt. **Is Garbage Collection Overhead Gone? Case study of F2FS on ZNS SSDs.** In the Proceedings of the 15th ACM Workshop on Hot Topics in Storage and File Systems (HotStorage), 2023.

NVMW'23 [W2] **Huaicheng Li**, Daniel S. Berger, Lisa Hsu, Daniel Ernst, Pantea Zardoshti, Stanko Novakovic, Monish Shah, Samir Rajadnya, Scott Lee, Ishwar Agarwal, Mark D. Hill, Marcus Fontoura, Ricardo Bianchini. **Pond: The Case of CXL Memory Pooling for Cloud Datacenters.** In the 14th Annual Non-Volatile Memories Workshop (NVMW), 2023.

Dissertation

[D1] Ph.D. Thesis, **Evolving Storage Stack for Predictability and Efficiency.** University of Chicago. 2020

[D2] M.S. Thesis, **FEMU: Fast, Accurate and Extensible Flash Emulator.** University of Chicago. 2018

Work Experience

Research Internships at Industrial Labs

Microsoft Research (Redmond), *Systems Research Group* Summer 2020
Research Intern working on resource disaggregation for datacenter deployment [ASPLOS'23]

Microsoft Research (Redmond), *Database Group* Summer 2019
Research Intern working on programmable storage

Microsoft Research (Redmond), *Systems Research Group* Summer 2018
Research Intern working on offloading cloud storage stack to ARM SoCs [ASPLOS'20]

NetApp, *Advanced Technology Group (ATG)* Spring 2020
Research Intern working on new file system designs for emerging storage hardware

Research Experience at Universities

Carnegie Mellon University, *Parallel Data Lab (PDL)* 2020–2022
Postdoctoral Researcher collaborating with Gregory R. Ganger, George Amvrosiadis, David G. Andersen and CMU students on new storage and memory technologies [ASPLOS'23]

University of Chicago, *Systems Group* 2015–2020
Graduate Student Researcher working on Operating and Storage Systems research [SOSP'21, ASPLOS'20, FAST'18, SOSP'17, FAST'17]

Wuhan University, *Cloud Computing Lab* 2012–2015
Research Assistant working on I/O virtualization and cloud resource scheduling

Engineering Internship in Industry

Tencent (Shenzhen) Summer 2012
Undergraduate Intern working on cluster resource monitoring and kernel optimization

Service

Program Committee (PC)

ASPLOS'25: The 30th ACM Intl' Conf. on Architectural Support for PL and OS	2025
ASPLOS'24: The 29th ACM Intl' Conf. on Architectural Support for PL and OS	2024
NSDI'24: The 21st USENIX Symposium on Networked Systems Design and Implementation	2024
CCGRID'24: The 24th IEEE/ACM International Symposium on Cluster, Cloud and Internet Computing	2024
ASPLOS'23: The 28th ACM Intl' Conf. on Architectural Support for PL and OS (*Fall cycle*)	2023
SySDW'23: Doctoral Workshop at SOSP 2023	2023
SYSTOR'23: The 16th ACM International Systems and Storage Conference	2023
NVMW'23: The 14th Annual Non-Volatile Memories Workshop	2023
APSys'21: The 12th ACM SIGOPS Asia-Pacific Workshop on Systems	2021
SYSTOR'21: The 14th ACM International Systems and Storage Conference	2021

External Review Committee (ERC)

ISCA'24: The 51st International Symposium on Computer Architecture (ISCA)	2024
ASPLOS'23: The 28th ACM Intl' Conf. on Architectural Support for PL and OS (*Spr/Smr cycles*)	2023

Organization Committee

FAST'24 Artifact Evaluation Committee, Co-chair	2024

Shadow Program Committee

EuroSys'20: The 15th European Conference on Computer Systems	2020
EuroSys'18: The 13th European Conference on Computer Systems	2018

Journal Reviewer

IEEE Micro	2022
TC: IEEE Transactions on Computers	2019, 2020, 2022, 2023
TOCS: ACM Transactions on Computer Systems	2019, 2023
TPDS: IEEE Transactions on Parallel and Distributed Systems	2019
JPDC: Journal of Parallel and Distributed Computing	2019, 2021
TOS: ACM Transactions on Storage	2018, 2023
CACM: Communications of the ACM	2018

External Reviewer

FAST'19: The 17th USENIX Conference on File and Storage Technologies	2019
ATC'18: The 2018 USENIX Annual Technical Conference	2018

Secondary/Sub-reviewer

FAST'20: The 18th USENIX Conference on File and Storage Technologies	2020
SOSP'19: The 27th ACM Symposium on Operating Systems Principles	2019
ASPLOS'19: The 24th ACM Intl' Conf. on Architectural Support for PL and OS	2019
FAST'18: The 16th USENIX Conference on File and Storage Technologies	2018

Artifact Evaluation Committee (AEC)

SOSP'21: The 28th ACM Symposium on Operating Systems Principles	2021

Departmental Service

Graduate Admission Committee, CS Dept, Virginia Tech	2023
Graduate Student Ministry - Minister for Faculty Hiring, CS Dept, University of Chicago	2019

Other Activities

VT CS Graduate Admission Committee	2024
VTURCS Research Symposium Judge	2023
ASPLOS'23 Mentoring Program	2023
Session Chair: ASPLOS'23 ("Storage Session")	2023
SOSP'21 Mentoring Program	2021
Session Chair: SYSTOR'21 ("Storage Session")	2021
USENIX HotStorage'20 Program Committee Meeting Scribe	2020
Chameleon Cloud Testbed Student Ambassador	2020
USENIX ATC'18 Program Committee Meeting Scribe	2018

Talks

Pond: CXL-based Memory Pooling for Cloud Platforms

Conference Talk, ASPLOS'23, Vancouver, BC, Canada	2023
Conference Talk, NVMW'23, San Diego, CA, USA	2023
Cornell Networked System Seminar	2023

Towards Predictable and Efficient Datacenter Storage

Invited Talk, Intel/VMware Crossroads 3D-FPGA Academic Research Center	2022

IODA: Host/Device Co-Design for Strong Predictability Contract on Modern Flash Storage

Conference Talk, SOSP'21, Online	2021
Parallel Data Lab (PDL) Seminar, Carnegie Mellon University, PA, USA	2021

Towards Hardware-based Memory Disaggregation

Invited Talk, Microsoft Research (Redmond), WA, USA 2020

NVMeFS: SmartNIC-centric File System Offloading
Invited Talk, NetApp, CA, USA 2020

Evolving Storage Stack for Predictability and Efficiency
Invited Talk, University of Illinois at Urbana-Champaign, IL, USA 2021
Ph.D. Thesis Defense, University of Chicago, IL, USA 2020
Invited Talk, Carnegie Mellon University, PA, USA 2020
Invited Talk, Microsoft Research - Cambridge, UK 2020
Invited Talk, University of California - Berkeley, CA, USA 2020
Invited Talk, University of Wisconsin - Madison, WI, USA 2020

LeapIO: Efficient and Portable Virtual NVMe Storage on ARM SoCs
Invited Talk, CAS ICT Young Scholar Forum, Beijing, China 2020
Conference Talk, ASPLOS'20, Lausanne, Switzerland 2020
Ph.D. Thesis Proposal, University of Chicago, IL, USA 2019
Invited Talk, Microsoft Research (Redmond), WA, USA 2018

The CASE of FEMU: Cheap, Accurate, Scalable and Extensible Flash Emulator
Master Thesis Defense, University of Chicago, IL, USA 2018
Conference Talk, FAST'18, Oakland, CA, USA 2018

Teaching

Instructor
CS5204: Operating Systems (Fall 2023), Virginia Tech 2023
CS6204: Advanced Operating Systems (Spring 2023), Virginia Tech 2023
CS3214: Computer Systems (Fall 2022), Virginia Tech 2022

Co-Instructor
18-746: Storage Systems (Fall 2021), Carnegie Mellon University 2021
Course website: https://course.ece.cmu.edu/ece746/index.html
↪ Co-teaching with Gregory R. Ganger and George Amvrosiadis, ∼100 students (BS/MS/PhD)
↪ Designing and giving lectures on storage management, file systems, etc.
↪ End-to-end class administration: course website, weekly TA meetings, designing quizzes, grading, etc.

Guest Lecturer
18-746: Storage Systems (Fall 2020), Carnegie Mellon University 2020
↪ Topic: "Ins and Outs of Storage Offloading using ARM SoCs"

Teaching Assistant
CMSC 230: Operating Systems, University of Chicago, TA 2015, 2018, 2019
↪ Hosting lab sessions about Pintos projects (tutorials/lectures, office hours, grading, etc.)
Computer Organization and Design, Wuhan University, TA 2014

Students

Current PhD Students (*indicates co-advised students)
1. Shoaib Asif Qazi 2023–Present
2. Inho Song* (with Sam Noh) 2023–Present
3. Jinshu Liu 2022–Present
4. Hamid Hadian 2022–Present
5. Yuze Li* (primary advisor: Ali Butt) 2021–Present
6. Yuyue Wang* (at UCLA) 2021–Present
7. Sumit Kumar Monga 2019–Present

Current MS Students
8. Subhalakshmi Selvanathan (ECE) 2022–Present

Current Undergraduate Students

9. Hongtao Zhang (VT) 2023–Present
10. Hansen Idden (ITS) 2022–Present
11. Fauzhan Wahyudi (ITS) 2022–Present
12. Muhammad Daffa Al Fahreza (Udinus) 2022–Present
13. Muhammad Akmal Arifin (ITB) 2022–Present

Alumni

14. Hanchen Xu (UCLA, intern) Summer 2023
15. Edward Halim, (BS → CS PhD student at University of Wisconsin - Madison) 2022–2023
16. Sumanth Rao (CMU MCDS MS → Snowflake) 2022–2023
17. Jiuzhi Yu (CMU MCDS MS) 2022–2023
18. Zixu Chen (CMU MCDS MS → Google) 2021–2022
19. Aditya Shetty (CMU MCDS MS → Google) 2021–2022
20. Fadhil I. Kurnia (ITB BS → CS PhD student at UMass - Amherst, co-author of [J2]) 2018–2020
21. Martin L. Putra (ITB BS → CS PhD student at University of Chicago, co-author of [C4, J2]) 2018–2020
22. Ronald Shi (UChicago BS/MS → Meta, co-author of [C4, J2]) 2018–2019

Student Awards/Recognitions

23. Jiuzhi Yu, Best Science Award (for CMU MCDS Capstone Project on Programmable Storage) 2022
24. Sumanth Rao, Best Science Award (for CMU MCDS Capstone Project on Programmable Storage) 2022

PhD Thesis Committee

25. Sumit Kumar Monga, VT ECE, **co-chair**
26. Yuze Li, VT CS

MS Thesis Committee

27. Subhalakshmi Ramakrishnapuram Selvanathan, VT ECE MS, **co-chair**
28. Sruthi Ayaluru Venkata Krishnan, VT ECE MS

Grants

1. "Rethinking System Stack for the Load-Store I/O Era" 2024–2029
   Huaicheng Li (PI). NSF CNS-2339901. ∼$677K
2. "Converged Memory and Storage Systems" 2024–2025
   Huaicheng Li (PI). Samsung. $250K
3. "A Cross-stack Approach to Reduce Memory Carbon for Cloud Data Centers" 2023–2026
   Huaicheng Li (Co-PI). NSF CNS-2312785. $1M
4. "Near-data Processing for Machine Learning Workloads Acceleration" 2023–2024
   Huaicheng Li (PI). 4-VA. $30K.
5. "CXL for Reduced Memory Management Tax" 2023–2024
   Huaicheng Li (PI). Samsung. ∼$270K.
6. "Disaggregation and Offloading for Improved System Efficiency" 2023–2024
   Huaicheng Li (PI). The Indonesian Ministry of Education, Culture, Research and Technology. $30K.
7. "Characterization Driven Data Placement Optimizations for CXL Memory" 2023–2024
   Huaicheng Li (PI). Samsung. ∼$110K (*research gift*).
8. "Enhancing Storage Stack Design for the Computational Storage Era" 2023–2024
   Huaicheng Li (Co-PI, **major proposal writer**). Samsung. ∼$120K.
9. New Faculty Mentoring Grant 2023–2024
   Huaicheng Li (PI). Virginia Tech. $1.5K.
10. "Fortified Computational Storage Stack for Efficient Application Offloading" 2022–2023
    Huaicheng Li (Co-PI, **major proposal writer**). Samsung. $110K.

Software

1. **RAZIN:** https://github.com/ZonedStorage/RAIZN-release 2023
   An array of independent zoned namespace SSDs built on top of a virtual zone interface.

2. **Pond:** https://github.com/vtess/Pond 2022
   A CXL memory emulator utilizing zero-core NUMA nodes with benchmarking results for more than 100 workloads.
3. **Queenie:** https://github.com/ucare-uchicago/Queenie 2022
   A user-level tool for extracting SSD internal properties.
4. **IODA:** https://github.com/huaicheng/IODA 2021
   A host/device co-designed flash array for strong deterministic performance.
5. **LeapIO:** https://github.com/huaicheng/LeapIO 2020
   A cost-efficient cloud storage stack design that has been deployed in Microsoft datacenters.
6. **FEMU:** https://github.com/vtess/FEMU 2018
   A popular storage research platform widely used by top venue papers at ASPLOS, FAST, OSDI, and SOSP, etc.
7. **MittSSD:** https://github.com/ucare-uchicago/mittssd 2018
   An OS design with millisecond service level agreement interface.
8. **ttFlash**: https://github.com/ucare-uchicago/tinyTailFlash 2017
   An SSD architecture design eliminating garbage collection overhead for tiny-tail latencies.
9. Linux Kernel Contributor: Linux Open-Channel SSD Subsystem - pblk (120☆) 2017

Media Coverage

Pond: CXL-Based Memory Pooling Systems for Cloud Platforms [ASPLOS'23]

Software Engineering, https://semiengineering.com 2023
The Next Platform, https://nextplatform.com 2022
Semi Analysis, https://semianalysis.com 2022
Tech Powerup, https://www.techpowerup.com 2022
Screen Hacker, https://www.screenhacker.com 2022

Fail-Slow at Scale [FAST'18]

The Morning Paper, https://blog.acolyer.org, search "fail slow at scale" 2018
ZDNet, https://www.zdnet.com/article/how-clouds-fail-slow 2018
Hacker News, https://news.ycombinator.com/item?id=16463714 2018