Yang Zhou

yangzhou1997.github.io
yangzhou@g.harvard.edu ◇ +1 617 599 8532
150 Western Ave, SEC 4.429, Allston, MA 02134, USA

Research Interests

Networked systems, operating systems, distributed systems, networking stacks, and network telemetry.

Education

Harvard University, Cambridge, MA, USA
Ph.D. in Computer Science *(Expected) June 2024*
M.S. in Computer Science *November 2021*
Thesis title: Codesigning Networking Stacks and Datacenter Applications for High Efficiency and Evolvability
Advisors: Minlan Yu and James Mickens

Peking University, Beijing, China
B.S. in Computer Science *July 2018*
Thesis title: Towards Faster and More Accurate Data Stream Processing
Advisors: Tong Yang

Work Experience

Harvard University, Research Assistant *August 2018–Present*

- *Kernel offloads:* Designed eBPF-based kernel offloads for distributed system protocols including Paxos (Electrode [1]) and serializable transactions (DINT [12]) to reduce kernel networking stack overhead. Implemented and evaluated atop unmodified Linux OSes, and achieved kernel-bypass-like throughput and latency.
- *μs-scale RPCs:* Designed an efficient inter-server load balancing scheme for μs-scale RPCs to achieve low tail latency and high goodput (Mew [11]). Implemented and evaluated for both kernel-bypass and kernel-based networking stacks.
- *SmartNIC architecture:* Designed and prototyped SGX-like trusted execution environments for network functions in SmartNICs under multi-tenant cloud environments (S-NIC [13]).

Google NetInfra Group and System Research Group, Student Researcher *June 2021–May 2023*

- *Far memory:* Designed an efficient far memory system that leverages erasure-coding, remote memory compaction, one-sided RMAs, and offloadable parity calculations to achieve fast, storage-efficient fault tolerance (Carbink [2]). Implemented and evaluated using production networking stack.
- *Distributed runtime:* Designed an efficient fault-tolerant distributed runtime based on tasks and actors by leveraging the Chandy–Lamport consistent checkpointing algorithm and causal logging mechanism.
- *μs-scale RPCs:* Identified and motivated the inter-server scheduling problem for μs-scale RPCs (leading to Mew).

VMware Research, Research Intern *July 2020–September 2020*

- *Geo-distributed data analytics:* Applied traffic redundancy elimination (TRE) technique to accelerate geo-distributed data analytics and save WAN traffic cost. Implemented atop Alluxio, an in-memory data cache system for analytics.

Facebook, Research Collaborator *November 2019–May 2020*

- *Network telemetry:* Conducted extensive measurement and analysis on Facebook's network telemetry system. Identified the importance of being evolvable and handling changes. Proposed a change cube abstraction to systematically track changes, and an intent-based layering design to confine and track changes (PCAT [3]).

SenseTime, Software Engineering Intern *March 2018–May 2018*

- *Distributed storage:* Worked on Ceph storage setup, testing, maintenance, monitoring, and alerting.

Peking University, Research Assistant *April 2016–July 2018*

- *Network telemetry:* Designed and implemented novel probabilistic data structures (e.g., sketches and Bloom filters) to optimize the memory usage, speed, and accuracy of network telemetry tasks (Cold Filter [4], Elastic Sketch [5], Pyramid Sketch [8], and more [6][15][19]).

PUBLICATIONS

Total 769 citations till November 2024 based on Google Scholar.

Conference Publications

[1] **Yang Zhou**, Zezhou Wang, Sowmya Dharanipragada, and Minlan Yu.
Electrode: Accelerating Distributed Protocols with eBPF. [link]
USENIX NSDI 2023.

[2] **Yang Zhou**, Hassan Wassel, Sihang Liu, Jiaqi Gao, James Mickens, Minlan Yu, Chris Kennelly, Paul Turner, David Culler, Hank Levy, and Amin Vahdat.
Carbink: Fault-Tolerant Far Memory. [link]
USENIX OSDI 2022.

[3] **Yang Zhou**, Ying Zhang, Minlan Yu, Guangyu Wang, Dexter Cao, Eric Sung, and Starsky Wong.
Evolvable Network Telemetry at Facebook. [link]
USENIX NSDI 2022.

[4] **Yang Zhou**, Tong Yang, Jie Jiang, Bin Cui, Minlan Yu, Xiaoming Li, and Steve Uhlig.
Cold Filter: A Meta-Framework for Faster and More Accurate Stream. Processing [link]
ACM SIGMOD 2018.

[5] Tong Yang, Jie Jiang, Peng Liu, Qun Huang, Junzhi Gong, **Yang Zhou**, Rui Miao, Xiaoming Li, and Steve Uhlig.
Elastic Sketch: Adaptive and Fast Network-Wide Measurements. [link]
ACM SIGCOMM 2018.

[6] Omid Alipourfard, Masoud Moshref, **Yang Zhou**, Tong Yang, and Minlan Yu.
A Comparison of Performance and Accuracy of Measurement Algorithms in Software. [link]
ACM Symposium on SDN Research (SOSR) 2018.

[7] Xiangyang Gou, Chenxingyu Zhao, Tong Yang, Lei Zou, **Yang Zhou**, Yibo Yan, Xiaoming Li, and Bin Cui.
Single Hash: Use One Hash Function to Build Faster Hash Based Data Structures. [link]
IEEE International Conference on Big Data and Smart Computing (BigComp) 2018.

[8] Tong Yang, **Yang Zhou**, Hao Jin, Shigang Chen, and Xiaoming Li.
Pyramid Sketch: A Sketch Framework for Frequency Estimation of Data Streams. [link]
VLDB 2017.

[9] **Yang Zhou**, Peng Liu, Hao Jin, Tong Yang, Shoujiang Dang, and Xiaoming Li.
One Memory Access Sketch: A More Accurate and Faster Sketch for Per-Flow Measurement. [link]
IEEE Global Communications Conference (Globecom) 2017.

[10] Junzhi Gong, Tong Yang, **Yang Zhou**, Dongsheng Yang, Shigang Chen, Bin Cui, and Xiaoming Li.
ABC: A Practicable Sketch Framework for Non-Uniform Multisets. [link]
IEEE International Conference on Big Data (BigData) 2017.

Papers Under Reviews

[11] **Yang Zhou**, Hassan Wassel, James Mickens, Minlan Yu, and Amin Vahdat.
Mew: Efficient Inter-Server Load Balancing for Microsecond-Scale RPCs. [link]
September 2023.

[12] **Yang Zhou**, Xingyu Xiang, Matthew Kiley, Sowmya Dharanipragada, and Minlan Yu.
DINT: Fast In-Kernel Distributed Transactions with eBPF. [link]
September 2023.

[13] **Yang Zhou**, Mark Wilkening, James Mickens, and Minlan Yu.
SmartNIC Security Isolation in the Cloud with S-NIC. [link]
October 2023.

Workshop and Demo Publications

[14] **Yang Zhou**, Hao Jin, Peng Liu, Haowei Zhang, Tong Yang, and Xiaoming Li.
Accurate Per-Flow Measurement with Bloom Sketch. [link]

IEEE International Conference on Computer Communications Workshops (INFOCOM WKSHPS) 2018.

Journal Publications

[15] Zhuochen Fan, Gang Wen, Zhipeng Huang, **Yang Zhou**, Qiaobin Fu, Tong Yang, Alex X Liu, and Bin Cui.
On the Evolutionary of Bloom Filter False Positives - An Information Theoretical Approach to Optimizing Bloom Filter Parameters. [link]
IEEE Transactions on Knowledge & Data Engineering 2022.

[16] Yuanpeng Li, Xiang Yu, Yilong Yang, **Yang Zhou**, Tong Yang, Zhuo Ma, and Shigang Chen.
Pyramid Family: Generic Frameworks for Accurate and Fast Flow Size Measurement. [link]
IEEE/ACM Transactions on Networking 2021.

[17] Tong Yang, Jie Jiang, **Yang Zhou**, Long He, Jinyang Li, Bin Cui, Steve Uhlig, and Xiaoming Li.
Fast and Accurate Stream Processing by Filtering the Cold. [link]
The VLDB Journal 2019.

[18] Tong Yang, Jie Jiang, Peng Liu, Qun Huang, Junzhi Gong, **Yang Zhou**, Rui Miao, Xiaoming Li, and Steve Uhlig.
Adaptive Measurements Using One Elastic Sketch. [link]
IEEE/ACM Transactions on Networking 2019.

[19] **Yang Zhou**, Omid Alipourfard, Minlan Yu, and Tong Yang.
Accelerating Network Measurement in Software. [link]
ACM SIGCOMM Computer Communication Review 2018.

TALKS

- Electrode: Accelerating Distributed Protocols with eBPF
  Duke University, ACE Center for Evolvable Computing, Google, USENIX NSDI — *April 2023*
  Columbia University — *March 2023*
- Carbink: Fault-Tolerant Far Memory
  Cornell University — *November 2023*
  WORDS workshop — *November 2022*
  Microsoft Research Redmond, USENIX OSDI — *July 2022*
  Google — *March & June 2022*
- Evolvable Network Telemetry at Facebook
  USENIX NSDI — *April 2022*
  Boston University, Meta — *March 2022*
- Cold Filter: A Meta-Framework for Faster and More Accurate Stream Processing
  Harvard University — *October 2018*

MENTORING EXPERIENCE

- Matt Kiley, Harvard College undergraduate — *2023*
  Accelerating distributed transactions using eBPF and AF_XDP-based RPC systems.
- Yunxi Shen, Tsinghua University undergraduate — *2023*
  Resource-efficient job scheduling in data centers.
- Xingyu Xiang, Peking University undergraduate — *2023*
  Accelerating distributed transactions using eBPF.
- Zezhou Wang, Peking University undergraduate → University of Washington PhD — *2022*
  Accelerating Paxos using eBPF (NSDI 2023, [1]).

TEACHING EXPERIENCE

- **Guest Lecture** on far memory, CS294-252: Architectures and Systems for Warehouse-Scale Computers, UC Berkeley
  Nov 2023

- **Teaching Assistant** for Prof. Minlan Yu, CS145: Networking at Scale, Harvard University *Spring 2021*
- **Teaching Assistant** for Prof. Tong Yang, Algorithm Design and Analysis, Peking University *Fall 2018*

PATENTS

- **Yang Zhou**, Hassan Wassel, Minlan Yu, Hank Levy, David Culler, and Amin Vahdat. "Fault Tolerant Disaggregated Memory". Pending (US20230185666A1), filed by Google in December 2022.

ACADEMIC HONORS

- Google Ph.D. Fellowship in Systems and Networking *2022*
- Finalist, Meta Ph.D. Fellowship in Networking *2022*
- Graduate Fellowship, Harvard University *2018*
- Excellent Bachelor Thesis (10/327), School of EECS, Peking University *2018*
- New Academic Star Award (1/193), School of EECS, Peking University *2018*
- Arawana Scholarship (2/193), Peking University *2017*
- Pinyou Hudong Scholarship, School of EECS, Peking University *2016*
- May Fourth Scholarship, Peking University *2015*

PROFESSIONAL ACTIVITIES

- PC Member: ACM SIGCOMM Poster/Demo 2023, IEEE INFOCOM Workshop on Networking Algorithms 2020.
- Reviewer (Conferences): ACM SIGKDD 2023.
- Reviewer (Journals): ACM Transactions on Modeling and Performance Evaluation of Computing Systems, IEEE/ACM Transactions on Networking, IEEE Journal on Selected Areas in Communications.
- Panelist: "Getting started with systems research" at Students@Systems 2022.

REFERENCES

Prof. Minlan Yu
Department of Computer Science
Harvard University
150 Western Ave, SEC 4.415
Allston, MA 02134, USA
+1 617 495 3986
minlanyu@g.harvard.edu

Prof. James Mickens
Department of Computer Science
Harvard University
150 Western Ave, SEC 4.416
Allston, MA 02134, USA
+1 617 384 8132
mickens@seas.harvard.edu

Dr. Amin Vahdat
Google Fellow and Vice President of Engineering
Google LLC
1600 Amphitheatre Parkway
Mountain View, CA 94042, USA
+1 650 390 7073
vahdat@google.com

Prof. Adam Belay
MIT CSAIL
32 Vassar St, 32-G996
Cambridge, MA 02139, USA
+1 617 253 0004
abelay@mit.edu

Dr. Ying Zhang
Senior Engineering Manager
Meta Platforms, Inc.
1 Hacker Way
Menlo Park, CA 94025, USA
+1 408 250 9961
zhangying@meta.com