Github: PanJason
(+41) 779940083 / (+86) 13716519899
Homepage: panjason.github.io
pyyjason@gmail.com / yueyang.pan@epfl.ch

Yueyang Pan Ph.D. Candidate

Education

EDIC Computer and Communication Sciences, EPFL Lausanne, Switzerland
Ph.D in computer science 2022 - 2027 *(expected)*

- Advisor: Prof. Sanidhya Kashyap
- Research area: Operating Systems and Machine Learning Systems

EECS, Peking University Beijing, China
B.S. in Computer Science and Technology; Minor in Finance 2018 - 2022

- GPA: 3.83/4.00, Rank: 12/230.
- Graduate with *Summa cum laude*
- Turing Class Honor Program
- Advisor: Prof. Chenren Xu
- Research area: Computer Networks and 5G Wireless Communication

Research Interests

My research focuses on efficiently leveraging heterogeneous resources—from memory (RDMA, CXL) to computing (CPU, GPU). Spanning the OS kernel to LLM serving frameworks, my work aims to improve system performance by optimizing resource utilization at each layer across the software stack.

Publications

1. **Scalable Far Memory: Balancing Faults and Evictions**
   Yueyang Pan*, Yash Lala*, Musa Unal, Yujie Ren, Seung-seob Lee, Yizhou Shan, Abhishek Bhattacharjee, Anurag Khandelwal, Sanidhya Kashyap (*Equal Contribution)
   ACM Symposium on Operating Systems Principles (SOSP 2025)

2. **Tolerate It if You Cannot Reduce It: Handling Latency in Tiered Memory**
   Musa Unal, Vishal Gupta, **Yueyang Pan**, Yujie Ren, Sanidhya Kashyap
   ACM SIGOPS 20th Workshop on Hot Topics in Operating Systems (HotOS XX)

3. **Transparent Multicore Scaling of Single-Threaded Network Functions**
   Lei Yan, **Yueyang Pan**, Diyu Zhou, George Candea, Sanidhya Kashyap
   European Conference on Computer Systems (EuroSys 2024)

4. **Monarch: A Fuzzing Framework for Distributed File Systems**
   Tao Lyu, Liyi Zhang, Zhiyao Feng, **Yueyang Pan**, Yujie Ren, Meng Xu, Mathias Payer, Sanidhya Kashyap
   USENIX Annual Technical Conference (ATC 2024)

5. **Ship your Critical Section, Not Your Data: Enabling Transparent Delegation with TCLOCKS**
   Vishal Gupta, Kumar Kartikeya Dwivedi, Yugesh Kothari, **Yueyang Pan**, Diyu Zhou, Sanidhya Kashyap
   USENIX Symposium on Operating Systems Design and Implementation (OSDI 2023)

6. **The First 5G-LTE Comparative Study in Extreme Mobility**
   Yueyang Pan*, Ruihan Li*, Chenren Xu (*Equal Contribution)
   ACM on Measurement and Analysis of Computing Systems (SIGMETRICS 2022)

7. **Critique of "A parallel framework for constraint-based Bayesian network learning via Markov blanket discovery"**
   Jiaqi Si, Junyi Guo, Zhewen Hao, Wenyang He, Ruihan Li, **Yueyang Pan**, Zhenxin Fu, Chun Fan
   IEEE Transactions on Parallel and Distributed Systems (TPDS 2022)

8. **Critique of "MemXCT: Memory-Centric X-Ray CT Reconstruction With Massive Parallelization"**
   Zejia Fan, Yuchen Gu, Zhewen Hao, **Yueyang Pan**, Pengcheng Xu, Yuxuan Yan, Fangyuan Yang, Zhenxin Fu, Yun Liang
   IEEE Transactions on Parallel and Distributed Systems (TPDS 2021)

Experiences

Meta | London 2025.06 -
Mentor: Nikolay Beloborodov, Usama Arif and SJ Park

- Integrate kernel memory allocation information with Strobelight
- Augment DAMON with different filters in virtual address space
- Improve the performance of linux swap for memory offloading

Alibaba | Hangzhou, China 2022.05 - 2022.09
Mentor: Nie Hao and Junchen Guo

- Develop testing framework for 5G Cloud Solutions
- Test and analyze the performance of different solutions

PKU Super Computing Team | Beijing, China 2021.04 - 2022.06
Mentor: Zhenxin Fu, Yun Liang and Chun Fan

- Participate in ASC Student Supercomputer Challenge (ASC) 20-21, SC Student Cluster Competition (SCC) 20, 21

Projects

Operating Systems

Mage 2023.09 - 2025.07
Mage extends the feasibility of memory offloading. We find existing solutions have severe multi-core scalability issues, preventing many apps from being offloaded. Our system, Mage, adopts design principles, such as asynchronous decoupling, pipelined execution,and contention avoidance, to address the issues and realizes them on both Linux and LibOS

NFOS 2021.10 - 2023.11
NFOS is a combination of programming model, runtime, and profiler for developing scalable network functions (NFs) on multicore servers. NFOS insulates developers from writing concurrent NF code as it is bug-prone and hard to scale. We show that serial, stateful NFs run atop NFOS achieve scalability on par with hand-optimized counterparts in Cisco VPP.

TCLock 2021.10 - 2022.12
TCLock is a family of locking protocols using transparent delegation, where a lock waiter automatically encodes its critical section information on its stack and notifies the combiner (lock holder). The combiner executes the shipped critical section on the waiter's behalf using a lightweight context switch. TClock requires zero app modification with superior performance.

Machine Learning Systems

MegaLLM 2024.08 -

MegaLLM is a new decoupled architecture with a memory daemon to manage KV caches across storage tiers (including SSD, DRAM and VRAM), minimizing transmission costs in the cluster.

Open Source

SGLang 2025.01 -

Work on Multi-tiering KV Cache and MLLM. Bug fix and performance optimizations

vLLM 2024.07 - 2024.12

Work on Multi-tiering KV Cache

Teaching

Data-intensive Systems, CS300, EPFL	2025.02
Advanced Operating Systems, CS477, EPFL	2024.09
Data-intensive Systems, CS300, EPFL,	2024.02
Introduction to Operating Systems, CS323, EPFL	2023.09
Introduction to Operating Systems, CS323, EPFL	2022.09
Computer Networks (Honor Track), 04832250, Peking University	2021.09
Summer Reading Group, Turing Class, Peking University	2019.08

Awards and Honors

• **Teaching Assistant Award**, EDIC, EPFL	2024.12
• **Outstanding Graduates of Bejing**, Beijing	2022.06
• **Outstanding Graduates of Peking University**, Peking University	2022.06
• **China National Scholarship**, China	2021.09
• **Merit Student**, Peking Univeristy	2021.09
• **ASC Student Supercomputer Challenge 20-21 4th Place**, Worldwide	2021.05
• **John Hopcroft Scholarship**, Turing Class	2020.09
• **Schlumberger Scholarship**, Peking University	2019.10

Skills

Languages: Chinese, English, French (A2).

Programming: `C++/C`, `Python`, `Shell`, `CUDA`, `Linux`.