YIRONG (EFFY) WANG

PhD student in Computer Science

Personal website | effywang57@gmail.com
+1 339 545 6591 | github.com/Effygal
Boston, MA | Linkedin

RESEARCH INTERESTS

Analytical Modeling: Cache analysis, system simulations, I/O benchmarking, QoS modeling and optimization, data profiling.

Computer Systems: Distributed storage, CXL-enabled systems, systems for AI.

SKILLS

Programming: C, C++, GDB, Bash, Python;
System emulation: QEMU, gem5;
Formal specification & verification: NuSMV, SPIN, TLA+, Z3.

Math tools: Wolfram Mathematica, Matlab.

SUMMARY

PhD candidate in Computer Science focused on cache analysis, storage benchmarking, and QoS optimization. Skilled in system modeling, simulation, and programming, with ten years of Linux experience and four years in designing and evaluating distributed storage solutions.

EDUCATION

08/2021 – Present **PhD Student in Computer Science** — **Northeastern University, USA**
Advisor: Prof. Peter Desnoyers
Lab: Systems and Storage Lab, NEU Systems Research Group

12/2012 – 08/2014 **Master of Science** — **University of Southampton, UK**
Graduated with Merit
Distinction in Thesis

09/2008 – 06/2012 **Bachelor of Engineering** — **Tianjin Polytechnic University, China**
GPA: 82/100

CURRENT PROJECTS

Analytical Modeling

IRM is not enough: Inter-reference distance analysis of FIFO and CLOCK replacement **Under submission**
We present a two-dimensional analytical framework that models both item frequency and recency. Our work includes the first FIFO and CLOCK miss ratio in closed-form for general hyperexponential traffic based on average eviction time approximation, a novel proof of the convexity of LRU miss ratio curves under IRM, and identification of scenarios favoring CLOCK and FIFO based on item frequency skewness.

Benchmarking Toolkit

Configurable and cache-accurate trace generation for storage benchmarking **Under submission**
2DIO is a cache-accurate synthetic trace generator for block storage workloads, creating traces which replicate complex cache behavior seen in many real-world workloads. It combines a small, quantized description of recency behavior with a simple frequency-based model, allowing it to mimic the hit ratio curve of nearly any trace, while scaling the generated trace up or down in time and footprint. It is the first storage workload generator based on inter-reference distance (recency), and the first fully-synthetic trace generator able to create workloads with strongly non-convex miss ratio behaviors.

QoS Optimization

Optimize the resource allocation for multi-tenant virtual disks in cloud storage **Ongoing Project**
We address the challenge of efficient and realistic resource allocation in multi-tenant cloud storage systems through a novel linear programming model. Our approach optimizes traffic allocation across virtual and physical resources while respecting both service-level agreements and device constraints, utilizing VM hypervisor functionalities rather than a centralized controller. Implemented using GLPK, our solution demonstrates improved resource utilization, offering cloud providers a scalable, decentralized tool to enhance performance and fairness in multi-tenant storage systems.

PAST PROJECTS

Distributed System

Distributed object storage built on top of ZNSSDs **Github link**
ZoneLog is a distributed object storage that employs ZNSSD's Zone Append semantics and utilizes the write pointer as a weak version ID for consistency. We have built a toy testbed called FilelogKV on top of the conventional file system that emulates ZoneLog by issuing conventional file system calls and utilizing the logic block address as the write pointer to explore the system behaviours.

Computational Storage

ISC-enabled LSM-tree for Read Optimization **Past project**
The LSM-tree is not inherently optimized for efficient reading and space utilization, as its design involves trade-offs. This project aims to enhance LSM-tree read performance and space efficiency by leveraging Computational Storage Drives, which offload the parallel search computations from the host to storage.

Computer Vision

Error3DVis: Interactive visualization of 3D Geometry with Errors **Past project**

Error3DVis is an interactive visualization tool designed to assist 3D vision researchers in assessing their outcomes in 3D reconstruction. It displays the 3D mesh derived from different reconstruction methods, and provides error and semantic heat maps using color encoding.

EXPERIENCE

08/2021 – Present **Computer systems research assistant & teaching assistant** **Northeastern University, USA**

- Proved convexity of LRU miss ratio curves using Average Eviction Time approximation; derived FIFO and CLOCK cache miss ratios in closed-form for IRM and hyperexponential traffics (submitted to Sigmetrics '24).
- Designed a configurable and cache-accurate I/O generator for storage benchmarking, enabling trace generation with complex cache behaviors (submitted to FAST '25).
- Collaborated with a cross-functional team in BU to design and implement distributed storage systems for the MOC projects.
- Assisted in teaching the graduate-level Computer Systems course (CS5600) for three semesters, offering tutorials and guidance to students on system programming, GDB debugging, and fundamentals of operating systems.
- Model QoS into metrics for optimization in real-world distributed data centers (ongoing research).

`Analytical Modeling / System Simulations / Distributed Storage / File Systems / Operating Systems / System Programming`

11/2014 – 08/2021 **IT Administrator** **Yunnan University, China**

- Administrate user accounts, deploy software and maintain data backup/recovery processes.
- Manage university network infrastructure, servers, and security systems
- Develop and document IT policies, monitor system performance, and resolve technical issues.

`Sysadmin / Operating Systems / Linux`