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, high-fidelity workload synthesis, QoS modeling and optimization.

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

SKILLS

Programming: C, C++, GDB, Bash, Python;
System emulation: QEMU, gem5;
Generative AI: GANs, PyTorch, CUDA, Optuna;
Formal methods: NuSMV, SPIN, TLA+, Z3;

Math tools: Wolfram Mathematica, Matlab.

SUMMARY

PhD candidate in Computer Science specializing in cache modeling, high-fidelity workload synthesis, and QoS optimization. Experienced in analytical modeling, system simulation and programming, with hands-on expertise in developing Generative Adversarial Networks (GANs) and hyperparameter tuning. Has ten years of Linux experience and four years of designing and benchmarking 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

Inter-reference distance analysis of FIFO and CLOCK cache replacement — **Under submission**
We extend Che's approximation to analyze CLOCK cache performance under both request models and address three key properties: (1) IRM always yields well-behaved, convex LRU miss ratio curves; (2) with a fixed inter-arrival time distribution, FIFO performs worse when item popularity is skewed; conversely, (3) CLOCK performs better when item popularity is skewed. We present formal proofs for these properties. We propose a practical two-dimensional analytical model for real workload analysis, combining frequency-based workload decomposition with recency-based modeling for each substream. We demonstrate that this approach accurately predicts FIFO and CLOCK cache performance for real-world workloads from the CloudPhysics corpus.

Benchmarking Toolkit

Configurable and cache-accurate trace generation for storage benchmarking — **Under submission**
We address two observed patterns in real-world storage I/O traces: (1) their inter-arrival distance (IRD) distributions show multiple cliffs and plateaus in the short term and converge to heavy-tailed distributions in the long term; (2) the cliffs (or plateaus) in their LRU miss ratio curve correspond to the spikes (or holes) in their IRD distribution. Based on these observations, we introduce 2DIO, a two-dimensional trace generation framework that encodes cache behaviors through a compact, quantized representation of recency patterns (IRD), combined with an independent reference model that characterizes frequency. This approach requires minimal parameters yet accurately reproduces complex, non-convex LRU miss ratio curves observed in real workloads. Our evaluation compares 2DIO to state-of-the-art methods, including deep learning approaches, demonstrating its cache accuracy and low cost. 2DIO scales with trace length and footprint while providing flexible configuration to replicate or customize a wide range of LRU miss ratio behaviors.

QoS Optimization

Dynamic Hierarchical Throttling for VMware vSAN — **Ongoing Project**
This project focuses on resource allocation in VMware vSAN using a dynamic linear programming model to optimize traffic across virtual and physical resources while adhering to SLA constraints. Instead of relying on a centralized controller, it leverages VM hypervisor capabilities to reduce queue delays. The approach also implements multi-hierarchical token control to enhance performance and fairness.

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**

- Research in cache modeling; study properties of different cache policies under varying workload assumptions.
- Design high-fidelity workload synthesis methods for various storage benchmarking applications.
- Collaborate with the VMware vSAN team to develop QoS modeling for fair resource allocation and performance isolation.
- Collaborate with cross-functional teams at Boston University to architect and deploy distributed storage systems for the Massachusetts Open Cloud (MOC) initiative.
- Assist in teaching the Computer Systems class (CS5600) for three semesters, offering tutorials and guidance to students on system programming, GDB debugging, and fundamentals of operating systems.

`Analytical Modeling / Simulations / Distributed Storage / 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 / Linux`