Shuwen Sun

Khoury College of Computer Sciences
805 Columbus Ave
Boston, Massachusetts, 02115 U.S.A.

email: sun.shuw@northeastern.edu
URL: https://ccs.neu.edu/ jethrosun

Research Interests

My research focuses on understanding distributed system behaviors and building fast, efficient, and reliable systems. In my PhD work I adapted techniques from newer storage device capabilities to build new distributed storage systems. My work also touches semantics and consistency guarantees of distributed systems.

Prior to my current work, I have worked on research projects in datacenter networking and performance diagnosis for distributed systems.

Education

9/2018 – 2025 (Expected) ***Northeastern University***, *Boston, Massachusetts*
Ph.D. in Computer Science
Advisor: *Prof. Peter Desnoyers*

9/2015 – 5/2017 ***Boston University***, *Boston, Massachusetts*
M.S. in Computer Science

9/2011 – 5/2015 ***Sun Yat-sen University***, *Guangzhou, China*
B.Eng. in Software Engineering

Skill Sets

- **Programming languages:** C/C++, Rust, Python, Go, Bash, SQL
- **Frameworks** DPDK, SPDK, SGX, eBPF, OpenStack, Docker, OpenTracing/OpenTelemetry, Boost
- **Tools:** Mininet, NS3, Git, Wireshark, LaTeX, D3.js, meson

- **Knowledge (networking):** Kernel bypassing (DPDK), Network Function Virtualization, Datacenter networking, Cloud networking, Software-defined Networking
- **Knowledge (storage):** Distributed storage, SSD, Zoned Namespace SSD
- **Knowledge (systems):** Distributed systems, Storage systems, End-to-end request tracing, Cloud computing
- **Knowledge (programming):** Rust (async, future, tokio), C++ (Boost ASIO)

Selected SWE Summaries

01/2023-Present **ZStore a distributed object store**, *Developer & Maintainer*

- Developed and built a distributed object store from ground up with embedded HTTP Server using SPDK.
- Total LOC (C++): 10 k
- With additional 2 k C++ for testing, and 2 k of Bash for infrastructure.

05/2022-09/2022 **Google SWE and Research Intern**, *Developer*

- Interned in Google Global Networking (GGN) within the Bandwidth Enforcer (BWE) Team
- Total CL: 12

01/2019-04/2022 **Building Network Functions with NetBricks**, *Developer & Maintainer*

- Developed 4 network functions (RDR proxy for web browsing, BitTorrent downloader, TLS certificate validator, and video transcoder) and 11 coresident NFs in NetBricks with only safe Rust and external Rust crates.
- Total LOC (Rust): 10 k
- With additional 600 of Rust, 3 kof Python, and 2.5 k of Bash for infrastructure.

5/2017 – 5/2018 **End-to-end Request Tracing with OpenStack**, *Developer*

- Worked on bootstrapping end-to-end tracing strategies into OpenStack's profiling tool, OSprofiler. Also worked on exploring ways to combine OpenTracing (now OpenTelemetry) with OSprofiler.

5/2016 – 1/2017 **Hardware Isolation Layer**, *Core contributor*

- Worked on the development and maintenance of HIL, a low-level tool for reserving physical machines and connecting them via isolated networks. Submitted 8 issues and 6 pull requests.
- LOC (Python): 2 k

Selected Research Summaries

1/2022 – now

ZStore: A Fast, Efficient, and Strongly-Consistent Object Storage System with ZNS SSDs

Mentor: *Prof. Peter Desnoyers* ☒ *Prof. Ji-yong Shin*.

- ZStore is a high-performance distributed S3-compatible object storage system designed to efficiently utilize modern SSDs. Unlike traditional object storage optimized for HDDs, ZStore leverages NVMe-over-Fabric and Zoned NameSpace append commands to implement independent per-device shared logs, reducing SSD read/write operations per request. It uses RDMA for fast gateway-to-gateway consistency communication and ensures strong read-after-write consistency with minimal coordination. By separating fast and slow paths, ZStore provides single-key linearizability. Evaluations show that ZStore outperforms Ceph and Minio by an order of magnitude, approaching the limits of SSD performance even on modest hardware, making it well-suited for modern workloads like IoT, serverless computing, and archival storage.

9/2022 – 1/2022 **A case for IO efficiency as a research metric for storage systems**

Mentor: *Prof. Peter Desnoyers* ☒ *Prof. Orran Krieger*

- This paper highlights the importance of IO efficiency as a key metric for evaluating distributed and disaggregated storage systems, beyond traditional performance measures like bandwidth and latency. With the rise of cloud-based storage, understanding internal trade-offs has become crucial for developers and operators. The authors propose methodologies for measuring and analyzing IO efficiency, emphasizing write amplification as a critical factor that varies across workloads and architectures. By identifying points of IO amplification within the storage stack, they demonstrate how architectural choices impact efficiency and, in turn, influence conventional performance metrics. The paper argues for treating IO efficiency as a first-class metric, integrating it into performance monitoring and system design to optimize modern storage systems.

5/2017 – 7/2018 **Pythia**, *A Cross-layer Just-in-time Instrumentation Framework for Debugging Distributed Systems*

Mentors: *Dr. Raja Sambasivan* ☒ *Prof. Orran Krieger*.

- Performance diagnosis is notoriously difficult with today's diverse and complex distributed applications. It is always hard to know a priori *where*, within *what* stack level, and at *what* granularity to add instrumentation to help diagnose problems that may occur far in the future. We build Pythia which is comprised of an dynamic tracing instrumentation library that allows developers to dynamically enable and disable instrumentation points in the system, and a reconstruction, analysis and diagnosis pipeline to identify problematic areas in near real-time.

9/2016 – 5/2017 **Enabling Flexibility in Cloud Networks with FlexNet**

Mentors: *Prof. Orran Krieger* ☒ *Prof. Rodrigo Fonseca*.

- We observe that there are two intertwined issues in datacenter networking: protocols are not designed to co-exist with each other, and there is no uniform way for tenants (and their applications) to select different protocols best suited their needs. In FlexNet, we proposed an architecture that decouples datacenter network to enable scheduling network resources in a decentralized manner. We showed that FlexNet is sufficiently

flexible to better support various kinds of applications while enabling niche protocols comprising of individual network technologies.

Publication

C4. **ZStore: A Fast, Efficient, and Strongly-Consistent Object Storage System with ZNS SSDs**
Sun, Shuwen, Khor, Isaac, Shin, Ji-yong, and Desnoyers, Peter.
Under submission

C3. **A case for IO efficiency as a research metric for storage systems**
Sun, Shuwen, Khor, Isaac, Krieger, Orran, and Desnoyers, Peter.
Under submission

C2. **Endpoint-defined In-Network Functions**
Sun, Shuwen and Choffnes, David.
Under submission

C1. **Toward Flexible Auditing for In-Network Functionality**
Sun, Shuwen and Choffnes, David.
CoNEXT-SW '22

Professional Experience

9/2018 – Present **Northeastern University**, *Graduate Research Assistant*
- Advisor: *Prof. Peter Desnoyers*

5/2023 – 8/2022 **Google**, *Research Intern*
- Mentor: *Vasileios Pappas*
- Ph.D. SWE intern in Google Global Networking
- CLs: 8 submitted, 12 in total

6/2022 – 9/2022 **ThousandEyes (Part of Cisco)**, *Research Intern*
- Mentor: *Dr. Arash Molavi Kakhki* ⊠ *Dr. Abbas Razaghpanah*
- Worked on Internet measurement related to anomaly detetion in HTTP timing. Part of the Internet Research team.

1/2018 – 7/2018 **Hariri Institute for Computing**, *Staff Researcher/Engineer*
- Supervisors: *Prof. Orran Krieger* ⊠ *Dr. Raja Sambasivan*.
- Worked on developing a novel diagnosis framework built upon end-to-end request tracing on the cloud. I worked on OpenStack's tracing system OSprofiler and benchmarking tool Rally.

6/2017 – 12/2017 **Massachusetts Open Cloud**, *Research Intern*

- Supervisors: *Prof. Orran Krieger* ⊠ *Dr. Raja Sambasivan*.
- Worked as a research intern at Red Hat Engineering as a part of the Distributed-system Tracing team within the MOC Research group. I worked on adopting end-to-end request tracing techniques for performance diagnosis on the cloud.
- Host: *Jan Mark Holzer*

9/2016 – 5/2017 **Massachusetts Open Cloud**, *Research Assistant*

- Supervisors: *Prof. Orran Krieger* ⊠ *Prof. Rodrigo Fonseca*.
- Worked with Da Yu, Prof. Orran Krieger, Prof. Rodrigo Fonseca, Dr. Raja Sambasivan and Dr. Jason Hennessey on FlexNet, a novel datacenter networking design. I also served as a core contributor to the HIL project.

5/2016 – 8/2016 **Massachusetts Open Cloud**, *Research Intern*

- Supervisors: *Prof. Orran Krieger* ⊠ *Prof. Peter Desnoyers*.
- Worked with Prof. Orran Krieger, Prof. Peter Desnoyers, and Dr. Jason Hennessey on Hardware Isolation Layer project.

2/2014 – 6/2015 **Sun Yat-sen University**, *Research Assistant*

- Supervisor: *Prof. Wushao Wen*.
- Performed research on networking and cloud computing. I worked on exploring Software Defined Networking and applying it to network QoS. I also worked on developing and deploying a OpenStack based customized cloud environment solution.

Honors *and* Awards

Academic Honors:

1/2018 Northeastern University Graduate School Ph.D. Fellowship.
(Admitted to Ph.D. program in Computer Science at Northeastern University)

Travel Grant Awards:

3/2021 NSDI '21 Conference Student Grant, USENIX
2/2020 NSDI '20 Conference Student Grant, USENIX
8/2019 SIGCOMM '19 Conference Student Grant, NSF
8/2017 SIGCOMM '17 Conference Student Grant, NSF
7/2016 ATC '16 and HotCloud '16 Conference Student Grant, USENIX

Miscellaneous:

2/2022 Invited participants of Google Network Research Summit
9/2014 Second-class Scholarship for Outstanding Students at Sun Yat-sen University
(Top 10%).

4/2011 Recipient of independent recruitment for Sun Yat-sen University in 2011 (Top 6%, roughly 660 of 11,000 in China).

Talks *and* Posters

Talks:

2/2025 "ZStore: A Fast, Efficient, and Strongly-Consistent Object Storage System with ZNS SSDs" **Shuwen Sun**. Talk at *Khoury Software Day 2025*, Boston, MA

2/2025 "ZStore: A Fast, Efficient, and Strongly-Consistent Object Storage System with ZNS SSDs" **Shuwen Sun**. Talk at *2025 New England Systems Day*, Boston, MA

1/2024 "A case for IO efficiency as a research metric for storage systems" **Shuwen Sun**. Talk at *2nd Northeastern System Day*, Boston, MA

1/2023 "Toward Flexible Auditing for In-Network Functionality" **Shuwen Sun**. Talk at *Student Workshop co-located with ACM CoNEXT 2022*, Rome, Italy

12/2022 "Toward Flexible Auditing for In-Network Functionality" **Shuwen Sun**. Talk at *1st Northeastern System Day*, Boston, MA

10/2017 "Pythia: A Just-in-Time Instrumentation Framework for Debugging Distributed Systems." Lily Sturmann, **Shuwen Jethro Sun**. Talk at *2017 MOC Annual Workshop.* Boston, MA

Posters:

3/2020 "How well does your network (function) function? Understanding Network Functions Under User-level Use Cases" **Shuwen Jethro Sun**, David Choffnes. Poster at *Khoury Ph.D. Open House.*

12/2017 "Pythia: A Cross-layer Just-in-Time Instrumentation Framework for Debugging Distributed Applications." Lily Sturmann, **Shuwen Jethro Sun**, Raja Sambasivan, Orran Krieger, Peter Portante. Poster at *IV New England Networking and Systems Day (NENS'17).* Boston, MA

10/2017 "Pythia: A Just-in-Time Instrumentation Framework for Debugging Distributed Systems." Lily Sturmann, **Shuwen Jethro Sun**, Rajul Kumar, Vladimir Pchelin, Orran Krieger, Peter Portante, Raja Sambasivan. Poster at *2017 MOC Annual Workshop.* Boston, MA

Teaching Experience

Spring 2024 **Head Teaching Assistant**,
CS 3650 Computer Systems, Northeastern
Course instructors: Peter Desnoyers and Cheng Tan.

Fall 2023 **Teaching Assistant**,
CS 5600 Computer Systems, Northeastern
Course instructors: Peter Desnoyers.

Fall 2021 **Teaching Assistant**,
CS 3700 Networks and Distributed Systems, Northeastern
Course instructors: David Choffnes, Sakib Miazi, Christo Wilson.

Professional Service

To Northeastern University ⊠ Khoury College:

9/2023-1/2024 • **Organizing committee ⊠ Program committee ⊠ Session chair** — 2nd Khoury Systems Day

1/2023 • **General Chair ⊠ Session Chair** — 1st Khoury Systems Day

7/2021 • **Co-organizer ⊠ panelist** — Workshop on Completing PhD Course Requirements at Khoury

3/2021 • **Moderator** — 2021 Ph.D. Open House Graduate Student Panel

1/2021 • **Ph.D. Open House Co-organizer** — 2021 Ph.D. Open House

1/2021 – 05/2022 • **Organizing member** — Khoury Graduate Students Association

12/2020 • **Ph.D. Admission Volunteer** — 2021 Ph.D. Admission

3/2020 • **Panelist** — 2020 Ph.D. Open House Graduate Student Panel

9/2019 • **Panelist** — 2019 Ph.D. Orientation

3/2019 • **Letter writer** — Khoury College Naming Ceremony Thank you Gift

3/2019 • **Panelist** — 2019 Ph.D. Open House Graduate Student Panel

Spring 2019 • **Co-organizer** — Systems and Networking Reading Group

To the Discipline:

2025 • **Reviewer** — IEEE Internet of Things Journal 2025

2024 • **Reviewer** — Peer-to-Peer Networking and Applications 2024

2024 • **Reviewer** — IEEE TIFS 2024

2021 • **External Reviewer** — CCS 2021

2020 • **Organizing Volunteer** — SIGCOMM 2020 Hallway Sessions

2020 • **External Reviewer** — IMC 2020

2019 • **External Reviewer** — NSDI 2020

Summer 2017 • **Layer 9 Scriber** — SIGCOMM 2017

Personal Trivia

Languages: English (professional proficiency), Mandarin (native)

Affiliations

- Khoury College of Computer Sciences, Northeastern University
- MOC Alliance

References

Peter Desnoyers, Associate Professor
Northeastern University
✉ pjd@ccs.neu.edu

Ji-yong Shin, Assistant Professor
Northeastern University
✉ j.shin@northeastern.edu

Orran Krieger, Research Professor
Boston University
✉ okrieg@bu.edu

Vasilis Pappas, Tech Lead
Google Global Networking
✉ vasilis@google.com