Yu Xingzi

edittriendl@sjtu.edu.cn | xingzyu.github.io

XingZYu | Institute of Scalable Computing and System, ISCS

Shanghai Jiao Tong University, Shanghai, China

Objective

I'm a Ph.D. candidate in Computer Science at Shanghai Jiao Tong University, supervised by Prof. **Zhengwei Qi**. My research focuses on Operating Systems, particularly in the areas of Virtualization and Disaggregated Memory.

Experience

- **Ant Group [🌐]** *Nov 2021 - April 2024*
  Research Intern Shanghai, China
  - Developed WASC, a secure container runtime for WebAssembly, offering function level sandboxing with system interface isolation.
  - Co-developed TS2WASM, an AoT compiler for TypeScript to WebAssembly, achieving 3.8× as many language features and up to 19× speed up.

Education

- **Shanghai Jiao Tong University** *Sep 2017 - June 2021*
  Bachelor in Computer Science (IEEE Honor Class) Shanghai, China
- **Shanghai Jiao Tong University** *Sep 2021 - Present*
  PhD in Computer Science Shanghai, China

Projects

- **HyperSwap (Anemoi): [A Hypervisor-based Disaggregated Memory System]** *May 2024 - Present*
  Tools: [KVM, QEMU, RDMA]
  - Developed a hypervisor-based disaggregated memory system that enables transparent deployment.
  - Implemented RDMA support for low-latency data transfer between memory and compute nodes.
  - Created a fault-tolerant mechanism based on Erasure codes to ensure high availability.
- **SplitLLM: [CPU-GPU Hybrid LLM Inference System]** *May 2024 - Present*
  Tools: [PyTorch, IPEX-LLM]
  - Developed a CPU-GPU hybrid inference system for large language models, optimizing resource utilization.
  - Implemented model partitioning techniques to balance workloads between CPU and GPU.
  - Created a dynamic scheduling algorithm to improve inference latency and throughput.
- **WASC: [WebAssembly Secure Container]** *Nov 2021 - April 2024*
  Tools: [WebAssembly, C, C++, Linux Kernel]
  - Developed a secure container runtime for WebAssembly, providing function-level sandboxing.
- **TS2WASM: [TypeScript to WebAssembly Compiler]** *Nov 2021 - April 2023*
  Tools: [TypeScript, WebAssembly, C++, LLVM]
  - Developed a TypeScript to WebAssembly AoT compiler, enabling development with TypeScript beyond traditional web environments.

Patents and Publications

C=Conference, J=Journal, P=Patent, S=In Submission, T=Thesis

[J.1] Xingzi Yu, et al. (2024). **Enhancing embedded systems development with TS^{-}**. *Automated Software Engineering*, Vol. 31, DOI: 10.1007/s10515-023-00404-x

[C.1] Xingguo Jia, **Xingzi Yu**, et al. (Year). **Rethinking Virtual Machines Live Migration for Memory Disaggregation**. In *2023 IEEE International Conference on Cluster Computing (CLUSTER)*, pp. 145-157. IEEE. October 31 - Nov. 3, 2023, Santa Fe, NM, USA. DOI: 10.1109/cluster52292.2023.00020

[S.1] Xingzi Yu, et al. (2025). **WaSC: Hardening WebAssembly Sandboxes via System Interface Decoupling**.

[S.2] Xingzi Yu, et al. (2025). **HyperSwap: Resilient Disaggregated Memory with Transparent Deployment**

[S.3] Xingzi Yu, et al. (2025). **SplitLLM: Adaptive Prefill-Decode Splitting of Large Language Models for Efficient Inference on Edge Devices**

[S.4] Xingzi Yu, et al. (2025). **Rethinking Virtual Machines Live Migration for Memory Disaggregation**. Manuscript submitted for publication in *IEEE Transactions on Parallel and Distributed Systems*

[J.2] Chen Chen, Haoyang Zhang, Kaicheng Guo, **Xingzi Yu**, et al. (2024). **Exploring Efficient Hardware Accelerator for Learning-Based Image Compression**. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 44, No. 6, pp. 2204-2217, DOI: 10.1109/TCAD.2024.3515856