



DISSERTATION DEFENSE



Yiming Qiu

Assisting Cloud System
Development with Automated
Insight Generation

Friday, November 22, 2024

4:00pm – 6:00pm

3725 Beyster

Hybrid – [Zoom](#)

ABSTRACT: As cloud computing revolutionizes the IT industry, the complexity of underlying cloud systems has been increasing exponentially. The advent of new hardware accelerators and software platforms has made it challenging for cloud users to master the ever-expanding development toolkits. Compounding the issue, the programming frameworks and internals of these new systems are highly heterogeneous, with vastly different performance characteristics, resource constraints, management principles, and reliability considerations. Consequently, it is becoming essential to minimize human effort when managing these new ecosystems. In this dissertation, we advocate for assisting cloud developers and operators by automatically generating system insights. Essentially, these insights bridge the gap between user intentions and system requirements, providing clarity on the outcomes of user actions on a system without the need for tedious trial-and-error processes. In this dissertation, I will demonstrate how we generate various types of insights for different cloud systems. Firstly, I explore performance optimization insights, which are critically needed as users attempt to offload legacy code from on-premise servers to emerging hardware accelerators like SmartNICs and DPUs. Secondly, I investigate safety compliance insights for the cloud networking stack, focusing on ensuring the correctness of system updates for the latest generation of runtime-programmable platforms. Lastly, I unearth infrastructure management insights for Infrastructure-as-Code (IaC) style cloud orchestration platforms, with the goal of mitigating the complexity caused by the fundamental cloud user/owner split. Throughout these projects, I leverage and extend techniques from a wide variety of disciplines, such as formal reasoning, software testing, machine learning, and their intersections. Our results demonstrate that we can generate useful insights across cloud data, control, and management planes, while unveiling an even larger insight generation and integration design space yet to be explored.

CHAIR: Prof. Ang Chen