

## Faculty Candidate Seminar -- Purdue Computes: Systems-Software

### Yang Zhou

Ph.D. Candidate, Computer Science Dept.  
Harvard University

**Tuesday, Feb. 27, 2024**  
**10:30 A.M. – 11:30 A.M.**  
**LWSN 3102 A/B**

### Network-Application Co-design for Efficient Datacenters

#### Abstract

Modern datacenters contain hundreds of thousands of servers and high-speed networks to run diverse applications. However, these datacenters suffer from low resource utilization and poor software performance that cannot be improved simply by relying on faster hardware. Because of these utilization and performance challenges, datacenters incur high operational costs, increased energy usage, and difficulty in handling growing application demands.

In this talk, I will focus on improving resource utilization and application performance through network-application co-design. I will first discuss how resource disaggregation, especially the far memory technique, is a promising way to improve memory utilization. However, prior research often lacks fault tolerance, a crucial requirement in datacenters. Subsequently, I will describe a fault-tolerant far memory system with network-efficient memory swapping and erasure coding, which requires far fewer network I/O operations than conventional wisdom, unlocking higher performance. I will then discuss how application-customized networking stacks can vastly improve the performance of network I/O-intensive distributed protocols such as consensus and transactions. The key insight is to safely offload protocol logic into kernel networking stacks to reduce kernel overhead. The resulting systems achieve the performance of kernel-bypass approaches but the security of kernel stacks.

#### Bio

Yang Zhou is a Ph.D. candidate in computer science at Harvard University, advised by Minlan Yu and James Mickens. His research is on systems and networking with a focus on improving resource utilization and software performance in large-scale datacenters. He takes a full-stack and cross-layer co-design approach to tackle practical systems problems. As part of his research, he has actively collaborated with companies such as Google and Meta. He is a recipient of a Google Ph.D. fellowship.