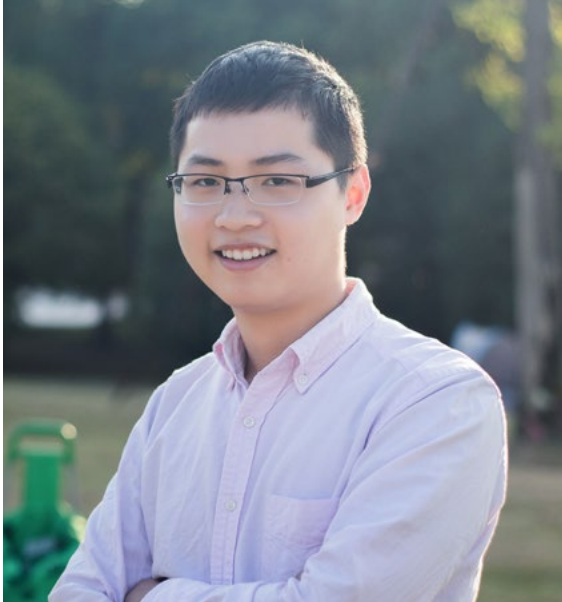




## DISSERTATION DEFENSE



# JIE YOU

## Toward Practical Application-Aware Big Data Systems

Tuesday, June 28, 2022

11:00am – 1:00pm

[Virtual](#)

**ABSTRACT:** In recent years, the growing number of Internet-of-Things (IoT) and other connected devices has created a surge in the amount of data being collected and analyzed. Data scientists gather useful insights from this data through data analytics and machine learning, all of which are executed on large-scale distributed systems known as big data infrastructures. They handle and abstract away low-level details such as resource allocation and task placements and provide simple APIs for application developers. While this decoupling between the infrastructure and applications makes developing applications easier, it also creates a gap between the system and applications. System designers optimize for generic performance metrics such as system-level throughput and latency while ignoring the application-level semantics, leading to suboptimal application performance. In this thesis, I explore the opportunities and means of co-optimizing application and infrastructure for a wide variety of big data applications. With the lessons learned from three projects, I discuss different design choices for bringing application-awareness into the infrastructure to improve the end-to-end performance and resource efficiency.

**CHAIR:** Prof. Mosharaf Chowdhury