Online services distribute and replicate state across geographically diverse data centers and direct user requests to the closest or least loaded site. While effectively ensuring low latency responses, this approach is at odds with maintaining cross-site consistency. In this talk, I will describe our three contributions to address this tension between consistency and performance. First, we propose RedBlue consistency, which enables blue operations to be fast (and eventually consistent) while the remaining red operations are strongly consistent (and slow). Second, to make use of fast operations whenever possible and only resort to strong consistency when needed, we identify conditions delineating when operations can be blue and must be red. Third, we introduce a method that increases the space of potential blue operations by breaking them into separate generator and shadow phases. I will also describe a coordination infrastructure called Gemini that offers RedBlue consistency, and report on our experience modifying applications to make use of it. Finally, I will describe our recent efforts on automating the choice of consistency levels for different application operations, and generalizing the notions and methods underlying RedBlue consistency.
Symposium on Operating Systems Principles, the flagship conference in computer systems, a special recognition award from MIT's Department of Electrical Engineering and Computer Science, and an ERC starting grant in 2012. In the last few years, he and his doctoral students have published their work in the top conferences of several areas, including OSDI, NSDI, EuroSys, FAST, USENIX Security, and ASPLOS.