Abstract

Although we are spending increasing amounts of time online, network performance and reliability problems hinder today’s applications and block the critical services we would like to run in the future. In this talk, I will present three projects to understand and improve the performance of cloud-based services. First, I will present work on reducing web latency by tailoring congestion response to the cloud setting. Our improvements reduced the web latency of Google clients by an average of 23%, and one of our techniques is now incorporated into Linux. Second, I will present techniques to map the serving infrastructure of cloud services. We used these techniques to uncover a massive expansion of Google infrastructure, influencing Microsoft’s approach to content delivery in turn. Third, I will present the PEERING testbed we developed to understand and test Internet routing. Traditionally, researchers are outsiders to the Internet ecosystem, passively measuring existing routes but unable to deploy proposed changes. PEERING connects directly to real networks around the world, allowing researchers to participate in the routing ecosystem, exchanging routes and traffic directly with the other networks. In describing these projects, I will highlight important features of the architectures of modern cloud-based services, and I will conclude by describing future work to use these features to realize the demanding applications of tomorrow.

Bio

Ethan Katz-Bassett is an Assistant Professor in the Department of Computer Science at USC. Prior to joining USC, he completed his PhD in 2012 at the University of Washington, advised by Tom Anderson and Arvind Krishnamurthy, and then worked at Google. He has earned awards including NSDI Best Paper Awards, an Applied Networking Research Award, an NSF CAREER Award, and Google Faculty Research Awards. His research goal is to make Internet services reliable and fast.