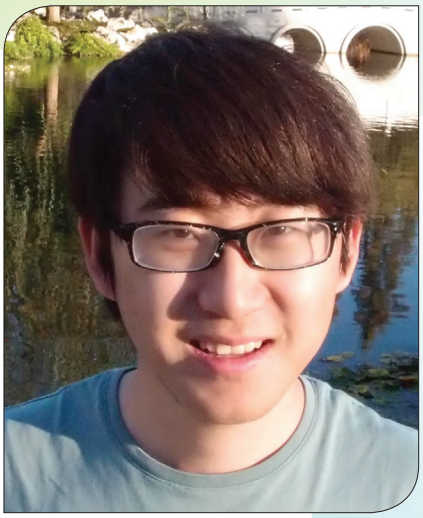
COLLEGE OF ENGINEERING
COMPUTER SCIENCE & ENGINEERING
UNIVERSITY OF MICHIGAN

CSE Dissertation Defense



Zhe Wu

Cost-Effective Support for Low Latency Cloud Storage

ABSTRACT: By offering data centers in several locations across the globe, cloud services enable web services to serve users from nearby data centers, thereby reducing user-perceived latencies. However, web services face many challenges in doing so due to the poor abstractions and lack of performance guarantees offered by cloud services. First, almost every cloud storage service offers an isolated pool of storage in each of its data centers, leaving replication across data centers to applications. Second, the limited low level interface offered by cloud storage makes it challenging for applications to implement any replication protocols. Third, popular public cloud storage services today experience high latency variance which can greatly degrade median application performance.

To address these problems, my thesis enables web services to combine the use of multiple cloud services leveraging the greater diversity in data center locations, different storage performance characteristics, and discrepancies in pricing. Spreading a web service's deployment across the data centers of multiple cloud providers however puts the onus of improving performance, reducing cost, and managing data replication on web services. Therefore, I have designed and implemented three systems that tackle these challenges. First, CosTLO judiciously combines several instantiations of the approach of issuing redundant requests and reduces high latency variance of cloud storage without relying on cloud providers to make any changes. Second, SPANStore greatly simplifies the task of geo-replicating data in a manner that cost-effectively satisfies any application's high-level latency, consistency, and fault-tolerance requirements. Lastly, CRIC enables strongly consistent access to geo-replicated data with latency and cost comparable to those achievable only if cloud storage offered a richer interface and had lower latency variance.

Chair: **Prof. Harsha Madhyastha**

Friday, July 14, 2017 2:00 – 4:00 pm 3725 Beyster Building