Lixin Gao
University of Massachusetts, Amherst
lgao@ecs.umass.edu

Topics of interest: 2. Network Support for the Cloud
4. Programming Models for the Cloud

Summary of Current Research:

My research has been focused on Internet routing, multimedia content delivery, data center networking, and task scheduling for parallel computation. My Ph.D. dissertation focused on task scheduling for parallel computation. Despite the fact that I have not pursed the topic further after my Ph.D. study, I continue to maintain the interest on the subject and have been paying close attention to the development of computing infrastructure in the cloud. Here I will focus on describing my research in the area of Internet routing and data center networking since they are the most closely related topics to Cloud Computing.

Routing in the Internet is achieved through exchanging network reachability information and making path selection decision through a distributed decision process. Hundreds of thousands of networks in the Internet make independent decision about path selection according to their own policies. Further, these policies do not have to be coordinated or simply the "shortest path" policy. As a result, it is possible that route might oscillate. Route oscillation can lead to packet loss and degraded network performance. My research has been focused on understanding the local policy that each network should use in order to ensure global routing stability. This is the first result that links economic model of the Internet with the performance of the underlying protocol. We are further able to infer the economic relationships between networks in the Internet, and the results are widely used by researchers in academics and industries, as well as network operators at major Internet backbone providers and content delivery networks (including AT&T, UUnet, Akamai, and Genuity).

Recently, I have been focused on network virtualization and data center networking. In network virtualization, we aim at providing programmability to virtualized networks without degrading the performance of each virtualized network. In data center networking, we design networks with commodity switches and PC hardware that sustain high network performance even under hardware failures.

Future Research Problems Planned to Pursue:

I plan to pursue research problems in the areas of programming models for the cloud, and network support for the cloud.

My research experience in parallel computation during my Ph.D. study will help me to pursue research in programming models for the cloud. Many of the existing parallel programming models have targeted at the computation intensive problems. With massive amount of data from Web, social media, and online social networks and myriad of applications such as machine learning and recommendation systems, it is imperative to design parallel programming model for massive datasets. While MapReduce is a programming framework proposed for data parallel applications, many well-known algorithms can not be implemented efficiently in MapReduce. For example, iterative algorithms or graph-based algorithms might not be programmed efficiently under the MapReduce programming framework. I plan to study programming models that facilitate the efficient execution of algorithms on massive datasets. In particular, we will focus on algorithms that can perform on-line queries in an efficient and robust manner.

Further, I will plan to address research issues in the data center networking and network virtualization in cloud. In data centers, commodity servers are interconnected through a hierarchy of routers. In the lowest level of the hierarchy, PCs are interconnected through Ethernet switches with hundreds of ports. The routers in the highest level of the hierarchy are expensive and high-end routers that are supposed to support communication between hundreds of thousands of servers. One of the key problems in data center networking design is to be able to scale to large number of servers with commodity hardware. Further the network has to be robust to failures in commodity servers and switches. For example, routing between servers has to survive failures in any of the component of the data center networks. I plan to pursue research in the design of data center networking. Further, I plan to pursue research on network virtualization in the cloud. The cloud is built to serve diverse customers who might have different needs. While computation and data storage are partitioned for individual customers, the network is shared among customers. Partitioning the network through network virtualization can provide isolation between cloud services. I plan to address the issues of network virtualization in the context of the cloud.