TCP Issues in Cloud Computing

Santosh Kulkarni

Graduate Student, Department of Computer Science and Software Engineering
Auburn University, Auburn, AL

Abstract

Data center and cloud computing, as closely coupling technologies, have evolved to become two of the most critical components of modern Internet. They have been used to build most of the large-scale Internet applications, such as Google search, Facebook, and Amazon. More importantly, they are the most prominent driving factors for innovation in the area of computing and networking.

Services of cloud computing are regularly hosted by tens of thousands of computers inside a data center, with significant aggregate bandwidth requirements. Due to cost and compatibility reasons, the communication fabrics of these centers are typically built using commodity Ethernet switches and routers. To meet the aggregate bandwidth requirements of data centers, evolving IEEE 802.3 standards have led to the development of 10 Gbps Ethernet networks. Such high speed Ethernet requires proportional scaling of TCP/IP processing so that, network intensive applications can ultimately benefit from the increased network bandwidth. While IP is expected to scale well in this context, TCP is known to have problems supporting very low latencies, high data rates, high availability and high robustness.

One such problem results in the catastrophic collapse in TCP's throughput when the number of servers sending synchronized data to a client increases past the ability of an Ethernet switch to buffer packets. In this talk we investigate the root causes for TCP's throughput collapse in data centers and examine a set of techniques that are effective in addressing this problem.

Bio

Santosh Kulkarni is a graduate student in the Department of Computer Science and Software Engineering, Auburn University. He received his Masters degree from the same department in May 2009 and is currently pursuing his doctoral degree under the guidance of his advisor, Dr. Prathima Agrawal. His active areas of research include Cloud Computing, Wireless LANs, Cooperative Networks and Distributed Computing.