

Paper Evaluation, Data Center TCP (DCTCP)

Jiani Jiang <jianij@kth.se>

1. Paper summary

Problems are proposed for TCP can not meet the increasing demands of apps suffered from bursty packet drops, Incast congestion and large queues. DCTCP is designed to solve the problems which combines ECN with a control scheme at the sources and extracts multi-bit feedback on congestion from the single bit stream of ECN marks. The marked packets are estimated by sources and used as a signal for congestion. It reacts in proportion to the extent of congestion which can reduce variance in sending rate and then can achieve large buffer headroom to fit bursts and attain low latency. The experiments illustrate that DCTCP reaches lower latency for short flows, high throughput for long flows and high burst tolerance for query flows compared with TCP.

2. Top 3 contributions

The new algorithm DCTCP proposed by the paper dramatically improves the performance of TCP. It achieves high burst tolerance, low latency and high throughput at the same time and makes it possible for data center networks to carry out mixing workloads and deliver better services. Meanwhile, it saves a large fraction of buffer space compared with TCP. It is backward compatible that it requires only 30 lines of code change to TCP, and setting of a single parameter on the switches.

3. Problems

1. DCTCP algorithm is not so widely supported. Data center servers may need to communicate with endpoints outside the data center to confirm if DCTCP is unsuitable. However, DCTCP algorithm provides no mechanism for negotiating functionality.
2. DCTCP and conventional TCP congestion control can't coexist well in a same network. DCTCP will use fractions of marked packets to slow down the rate reduction but TCP will suffer packet loss and quickly and drastically reduce cwnd.