xDCTCP: Extended DCTCP

Piyush Anil Nahar, Ajeet Kumar, Venkata Satya Sri Harsha Aduri
Computer Science and Engineering, Fall 2015

Introduction

Prologue
DCTCP (*SIGCOMM 2010*) solves problems specific to DCs like high latency and queue build-up on switches by reacting to congestion in proportion to the extent of congestion. But it treats both long and short flows alike however their requirements are different.

Goals
An improved protocol which improves upon DCTCP by reacting to congestion depending on the size of flow to reduce the latency of query and short flows while maintaining the high throughput for long flows.

Methodology

Key Idea: Short flows are more latency sensitive than Long flows and hence we reduce the window size of short flows by a lesser factor than long flows.

The estimate of fraction of packets marked with ECN is given by α based on the F (fraction of packets marked) and g (weight given to new samples)

$$\alpha = (1 - g) \times \alpha + g \times F$$

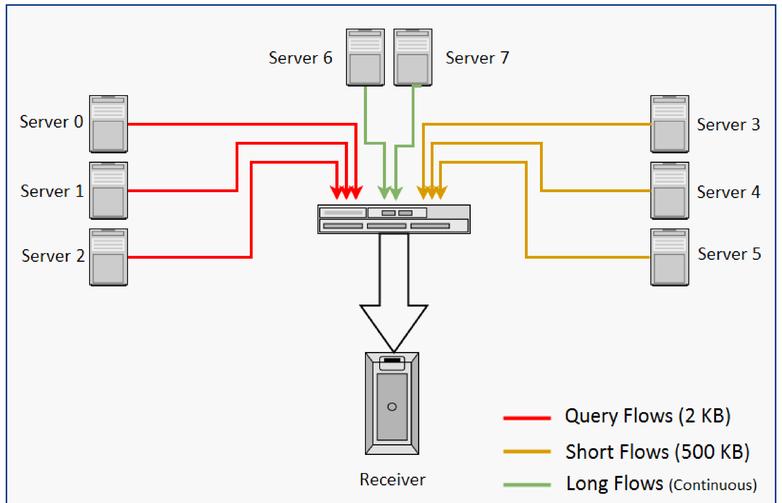
We modify congestion window for long flows and short flows differently based on a separate congestion control parameter for each of them.

$$cwnd = cwnd \times (1 - dctcp\_nl\_ \times \alpha)$$
$$cwnd = cwnd \times (1 - dctcp\_ns\_ \times \alpha)$$

We perform an analysis for several possible combinations of ***dctcp_ns_***, ***dctcp_nl_*** and ***g*** to find the optimal solution set which satisfies our design goal.

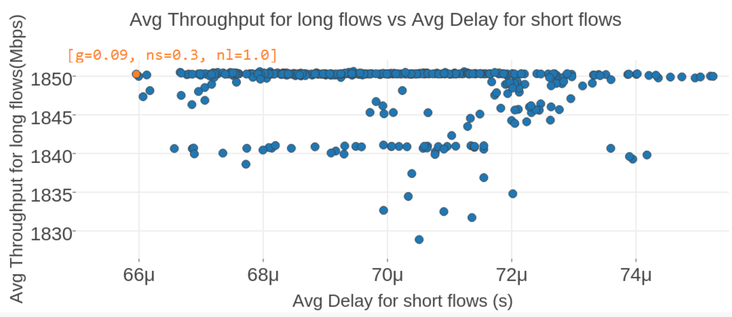
$$0.05 \leq g \leq 0.11 \; in \; steps \; of \; 0.005$$
$$0.1 \leq dctcp\_ns\_ \leq 0.5 \; in \; steps \; of \; 0.1$$
$$0.5 \leq dctcp\_nl\_ \leq 1 \; in \; steps \; of \; 0.1$$

Topology and Traffic Flows



Parameter	Value	Parameter	Value
Link Speed	10 Gbps	Simulation time	1s
Round Trip Time	100µs	ECN Marking Threshold	65 packets
Query Flow arrival interval	1ms	Switch buffer	250 packets
Short Flow arrival interval	10ms	Packet size	1460 bytes

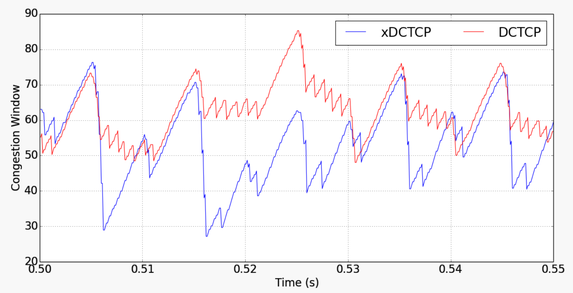
Results



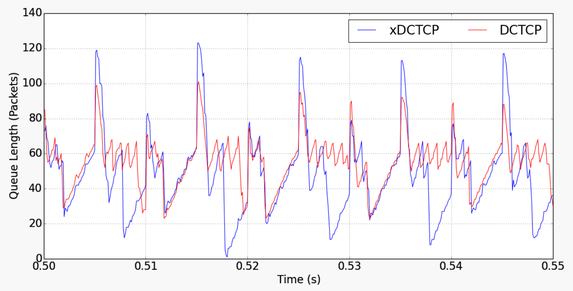
In the first stage, we compute the triplet [*g, ns, nl*] for which we get highest possible throughput with least average delay. We get *65.955µs delay* for short flows and *1850.26 Mbps* throughput for long flows at $g = 0.09$, $ns = 0.3$ and $nl = 1.0$

Results

We compare the congestion window and queue length variations with DCTCP.



- xDCTCP more aggressively cuts the congestion window for longer flows.



- The occasional spikes in queue length are a result of the burst traffic of the short flows occupying the queue
- Also, we obtain consistently similar values of *[g,ns,nl]* for various topology configurations and traffic patterns that satisfy our design goal.

Conclusion

- We have successfully reduced the average delay of short flows by 6.646% over DCTCP with a marginal 0.02% increase in throughput.
- We conclude that reducing the window size of longer flows by greater factor results in lower latency for short flows

Tools: NS2, Plotly, Matplotlib