Efficiency of BQL Congestion Control under High Bandwidth-Delay Product Network Conditions

Nikolai Mareev, Dmitry Kachan, Kirill Karpov, Dmytro Syzov and Eduard Siemens
Future Internet Lab Anhalt, Anhalt University of Applied Sciences, Bernburger Str. 57, 06366 Köthen, Germany,
{nikolai.mareev, dmitry.kachan, kirill.karpov, dmytro.syzov, eduard.siemens}@hs-anhalt.de

Keywords: Long Fat Networks, Transport Protocol, IP Networks, High-Speed Network, Congestion Control.

Abstract: BQL congestion control is aimed to utilize full available bottleneck bandwidth while keeping bottleneck buffer queue load on some low level to prevent it from producing avoidable additional delays or delay jitter. In this paper, an intermediate result of research in delay-based congestion control is presented. Using RMDT protocol we have evaluated its performance under high bandwidth delay product network conditions and compared it with TCP BBR using the iperf utility. High bottleneck bandwidth utilization in a wide area of delay/bandwidth/loss conditions have been reached. Some performance issues of BBR in some cases has also been observed and investigated.

1 INTRODUCTION

Congestion control algorithms take a significant role in the efficiency of data transport over IP networks.

In general, there are three main challenges for a modern congestion control algorithm: high bottleneck bandwidth utilization, resource sharing, and low influence on network buffers. The most common congestion control type in a network is loss-based congestion control. Using packet losses as a congestion indicator leads to performance degradation in lossy networks, and additional delays caused by the bottleneck queue load. Available buffer space of the bottleneck queue buffers significantly increased last time, what makes this additional delay significant. These consequences show the need of new congestion control algorithms with lower influence on the network and with non-congested packet loss tolerance.

One such solution is a BQL (Bottleneck Queue Level) congestion control developed in the course of the CloudBDT and BitBooster projects at the Future Internet Lab Anhalt. Mentioned projects operate with a Reliable Multi-Destination Transport protocol RMDT [1], [2]. It is a delay-based congestion control solution with packet loss tolerance and low influence on the network infrastructure.

The aim of this paper is to present intermediate results the actual advances in research on a delay-based congestion control. For this, a series of tests of the efficiency of data transport using the developed BQL algorithms have been performed.

Results of such a solution in high bandwidth delay product network conditions in comparison to TCP BBR [4] (Bottleneck Bandwidth and Round-trip propagation time) have been shown. TCP BBR - is besides BQL another modern congestion based congestion control solution with similar aims as BQL and it is the closest solution to the proposed algorithm.

The content of this paper is organized as follows: In section 2 a brief observation of data transport issues over high bandwidth delay product network is provided. Section 3 describes the testbed network. In section 4 test scenarios and test results are provided. Section 5 includes conclusions over experiments and further work.

2 RELATED WORK

The first test results of a BQL congestion control were presented in the 6th International Conference on Applied Innovations in IT [3], (ICAIIT 2018). The main idea of this solution is to use a modified PID (Proportional – integral – derivative) controller to keep link always slightly congested with the aim to reach full bottleneck bandwidth utilization. The most significant states of an algorithm in the current implementation are the Gain state (to quickly reach bottleneck bandwidth limit) and the Control state (to keep bottleneck slightly congested). In [3] first performance tests have been provided, which show a fair resource sharing capability, full bottleneck bandwidth utilization and the overall structure of the BQL algorithm.

In this paper, new results after some development period of the algorithm are presented. Hereby, the stability and efficiency of congestion control have been significantly increased. Transport delays (delay between action and reaction in terms of a controller) generally caused by RTT in the current version of the algorithm have now much less destructive influence on the performance of control, what allows to keep a necessary number of bytes in a bottleneck queue buffer more precisely.

Most changes during development were done in control state of an algorithm. The current version of RMDT allows using more accurate delay metrics of network congestion which lead to higher performance in long pipes. Paired modified PID controllers now can stabilize throughput on full utilized bottleneck bandwidth under extreme conditions of up to 1000 ms of RTT delay with nearly 200 KB of memory usage by network device queue buffer. It is a benefit in the context of usage in networks with high throughput and tiny buffers. Paired controllers provide more precise control in cases with network delay jitter what leads to higher performance in noisy networks.

Another modern congestion-based congestion control solution is TCP BBR which provides high bottleneck bandwidth utilization in a wide area of conditions while keeping buffer load on some low level. The mechanism of keeping high bandwidth utilization of this solution is a bandwidth probing what leads to RTT jitter and rate losses caused by congestion control which is increasing with the growth of network delay.

Performance degradation of TCP BBR during resource sharing in 1Gbps and 10Gbps links is presented in [5]. Here shown that small bottleneck buffers can lead to packet losses caused by congestion during resource sharing and unfair coexistence of TCP BBR with other congestion control, especially in cases with diverse flow round trip time. In [6] a cyclic performance drop of TCP BBR was observed. However, BBR shows higher performance in cellular networks [7] in comparison to other congestion control algorithms. In [8] a detailed analysis of TCP BBR algorithm behaviour is presented. In this work performance degradation of BBR in cases with shallow buffers caused by overestimating the bottleneck capacity has been observed. In these cases, BBR cannot recognize that the network is congested what leads to datarates higher than available bandwidth and so to massive packet losses.

It is worth to mention that in many other cases, BBR congestion control algorithm can reach high bottleneck bandwidth utilization along with keeping low mean bottleneck buffer load level and nearly fair coexistence with other flows.

3 TESTBED NETWORK

The testbed network topology for our investigations is presented in Figure 1.

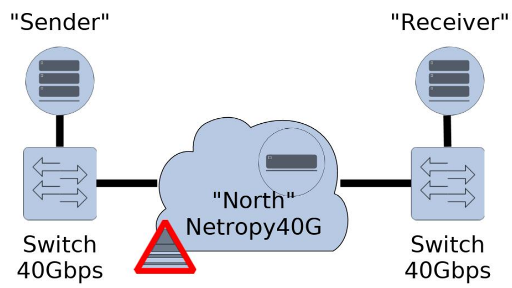

Figure 1: Test network setup.

WAN emulator Netropy 40G is the core element here, it can be used to emulate WAN links with up to 40 Gbps throughput and up to 1000s delay and to collect different statistics such as datarate and bottleneck buffer load level. Both servers run in Ubuntu 16.04 (kernel: GNU/Linux 4.15.0-45-lowlatency x86_64) and are equipped with Intel(R) Xeon(R) CPU E5-2643 v4 3.40GHz, 64GB of RAM and 40000baseSR4/Full supported link modes on Emulex Corporation OneConnect NIC.

The first bunch of test aimed to evaluate behavior of BQL under different round trip time conditions and comparing its performance with TCP BBR. The second bunch of tests is aimed to demonstrate the performance of BQL in a wide range of round trip time / packet loss rate / bottleneck bandwidth conditions. Mean datarate mentioned in these tests refers to the amount of transmitted data divided by time at the sender elapsed to transmit it. All tests have been performed in 40 GE Laboratory of Future Internet Lab Anhalt (FILA).

4 EXPERIMENTAL RESULTS

Results of the first bunch of tests are presented in Figures from 2 to 4. Each of these tests were evaluated over testbed network with next parameters of Netropy link: BBW (bottleneck bandwidth) = 1 Gbps; RTT = {0, 100, 500, 1000} ms; Queue buffer size = 80 MB, drop tail queuing algorithm. TCP BBR flows were executed with iperf utility.

On Figure 2 differences in behavior between BBR and BQL are shown.

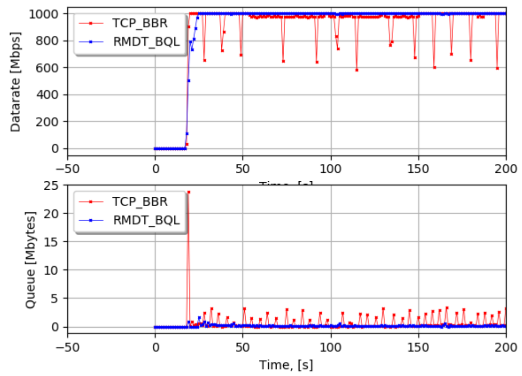

Figure 2: BBW 1 Gbps, base RTT 100 ms.

Both algorithms reach maximum available bandwidth. 75 Gigabytes were transmitted in approximately 10 min. Mean bottleneck buffer load levels during transmission were: 158 KB occupied by BQL and 705 KB occupied by BBR. Mean datarate during transmission were: 994.6 Mbps by BQL and 964.2 Mbps by BBR. The most significant difference between these two flows is buffer jitter. For BBR it can reach up to 4 MB while buffer jitter caused by BQL is less than 200 KB during the Control state period.

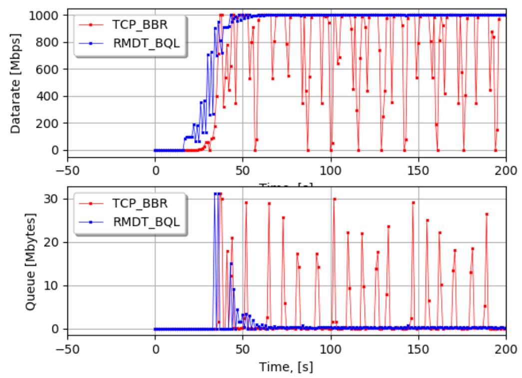

Figure 3: BBW 1 Gbps, base RTT 1000 ms.

Figure 3 demonstrates a more severe difference between these solutions: active bandwidth probing by BBR leads to significant rate decay during transmission and buffer jitter up to 30 MB. Mean bottleneck buffer load level during transmission was: 0.387 MB occupied by BQL and 3.633 MB occupied by BBR. Mean datarate during transmission was: 972.9 Mbps by BQL and 821.6 Mbps by BBR.

In Figure 4 mean datarates and buffer load level caused by these algorithms in different network delays are shown.

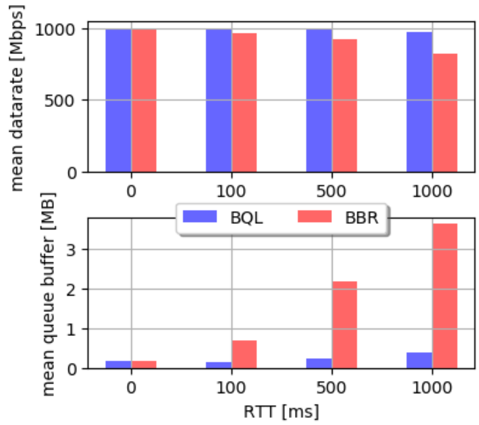

Figure 4: BBW 1 Gbps, 75 GB, summary results.

On Figures 5 and 6 a comparison between algorithms under packet loss conditions is presented. Both BBR and BQL do not use packet losses as congestion indicator what allows reaching high bandwidth even in presence of significant PLR (Packet Loss Rate). During 1 Gbps tests here in each case, a 60 Gb of data have been transmitted.

Figure 5 reveals the growing performance difference between BBR and BQL with increasing of RTT delay. Nevertheless, both algorithms provide high performance in such cases.

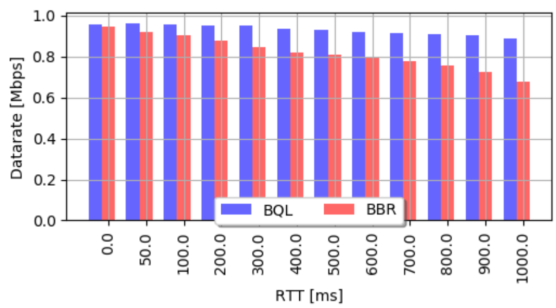

Figure 5: BBW 1 Gbps, PLR 1 %, 60 GB.

Figure 6 demonstrates tests bunch under 10 Gbps bottleneck bandwidth, 0.7 % packet loss rate and variety round trip time delays.

TCP stack of both server and a receiver was tuned up to its maximum but it turned out that it is not enough for such test conditions. It can be seen that RMDT under these conditions has a key advantage – a user-space protocol buffers and faster lost packets processing algorithm.

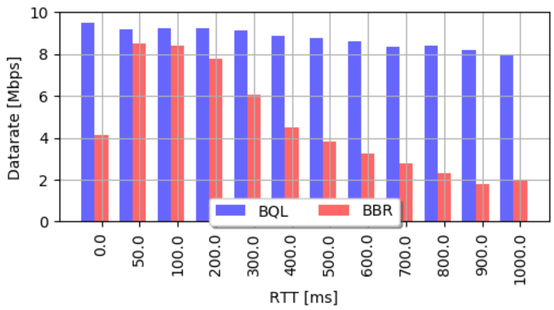

Figure 6: BBW 10 Gbps, PLR 0.7 %, 200 GB.

On 0 ms RTT case (in fact there present small network delay of appx. 150 µs) BBR showed significant performance drop. BQL as the congestion control algorithms in RMDT keeps on reaching high performance. During 10 Gbps tests in each case, a 200 GB of data has been transmitted.

5 CONCLUSIONS AND FURTHER WORK

In this article performance investigations of BQL in high bandwidth delay product network has been provided. With the main aims of BQL – to be scalable to different link cases, this solution can provide high bottleneck bandwidth utilization in wide conditions area. The raising RTTs does not have a significant influence on its control performance. Bottleneck buffer load level during all tests was kept on a low level, the mean value of bottleneck buffer load level during all 1 Gbps tests was nearly 250 KB. Packet losses do not have a significant effect on congestion control performance. Comparison with TCP BBR under the same conditions is also provided. This solution can provide high performance in many cases. However, performance degradation in high bandwidth delay product conditions was observed.

BQL congestion control algorithm is under active development. One of the main aims of the next work is an adjustable resource sharing algorithm - providing a mechanism of fair / low priority / aggressive coexistence of BQL with loss-based and delay-based common TCP congestion control algorithms. Boosting performance in wireless networks in common and in wi-fi networks in particular.

ACKNOWLEDGMENTS

This work has been funded by Volkswagen Foundation for trilateral partnership between scholars and scientists from Ukraine, Russia and Germany within the project CloudBDT: Algorithms and Methods for Big Data Transport in Cloud Environments.

REFERENCES

[1] E. Siemens, D. Syzov, D. Kachan High-speed UDP Data Transmission with Multithreading and Automatic Resource Allocation Proc. of: the 4th International Conference on Applied Innovations in IT, (ICAIIT 2016), pp. 51-56, Koethen, 2016.

[2] S. Maksymov, D. Kachan, E. Siemens Connection Establishment Algorithm for Multi-destination Protocol Proc. of: the 2nd International Conference on Applied Innovations in IT, (ICAIIT 2014), pp. 57-60, Koethen, 2014.

[3] N. Mareev, D. Kachan, K. Karpov, D. Syzov, E. Siemens, and Y. Babich, "Efficiency of a PID-based Congestion Control for High-speed IP-networks," Proc. of: the 6th International Conference on Applied Innovations in IT, (ICAIIT 2018), pp. 1-5, Koethen, 2018.

[4] N. Cardwell, Y. Cheng, C. S. Gunn, S. H. Yeganeh, and Van Jacobson, "BBR: congestion-based congestion control," Communications of the ACM, vol. 60, no. 2, pp. 58-66, Jan. 2017.

[5] M. Hock, R. Bless, and M. Zitterbart, "Experimental evaluation of BBR congestion control," in 2017 IEEE 25th International Conference on Network Protocols (ICNP), 2017, pp. 1-10.

[6] K. Miyazawa, K. Sasaki, N. Oda, and S. Yamaguchi, "Cyclic Performance Fluctuation of TCP BBR," in 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC), 2018, vol. 01, pp. 811-812.

[7] E. Atxutegi, F. Liberal, H. K. Haile, K. Grinnemo, A. Brunstrom, and A. Arvidsson, "On the Use of TCP BBR in Cellular Networks," IEEE Communications Magazine, vol. 56, no. 3, pp. 172–179, Mar, 2018.

[8] D. Scholz, B. Jaeger, L. Schwaighofer, D. Raumer, F. Geyer, and G. Carle, "Towards a Deeper Understanding of TCP BBR Congestion Control," IFIP Networking 2018, pp. 109-117, Zurich, Switzerland, May, 2018.