

Performance Evaluation of TCP BBRv2 Alpha for Wired Broadband, considering Buffer Sizes, Packet Loss Rates, RTTs, and Number of Flows

Elie Kfoury, Jose Gomez, Jorge Crichigno
University of South Carolina

<http://ce.sc.edu/cyberinfra>

ekfoury@email.sc.edu, gomezgaj@email.sc.edu, jcrichigno@cec.sc.edu

CI Engineering Lunch and Learn - Online
March 26, 2021

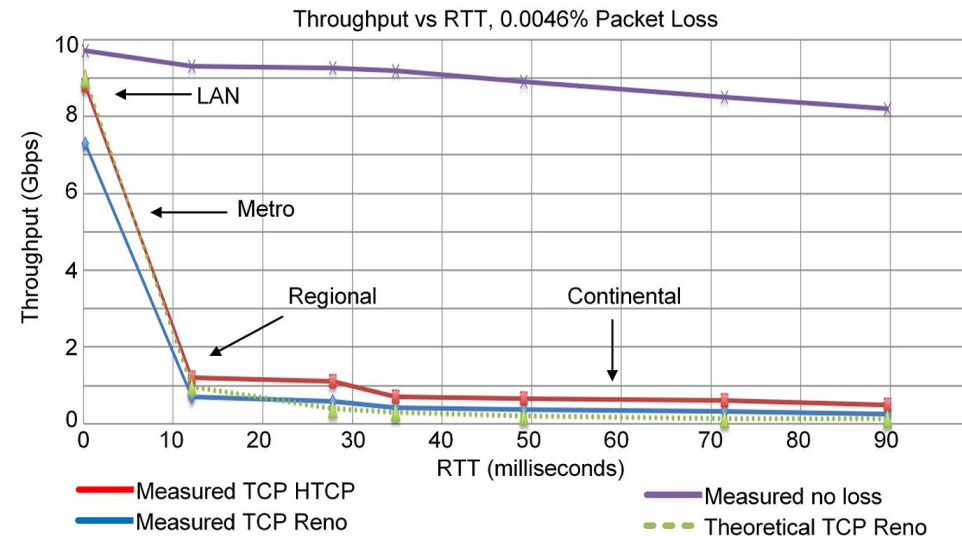
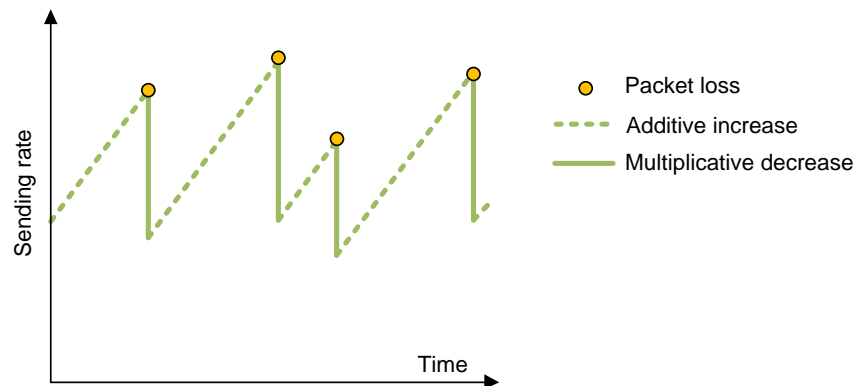
We want to recognize Neal Cardwell for his helpful comments and suggestions

Agenda

- TCP loss-based congestion control
- Bottleneck Bandwidth and Round-trip propagation time (BBR)
- BBRv1
- BBRv2 and design goals
- Experimental setup
- Results and evaluations
- Conclusions

TCP Loss-based Congestion Control (CC)

- The principles of loss-based CC were described in the 1980s¹
- Traditional CC algorithms follow the additive-increase multiplicative-decrease (AIMD) form of congestion control



1. V. Jacobson, M. Karels, Congestion avoidance and control, ACM SIGCOMM Computer Communication Review 18 (4) (1988).

BBRv1

- BBRv1 was published in 2016
- It produces high throughput in the presence of packet losses
- However, it shows some performance issues
 - Unfairness with loss-based CC (e.g., Reno, CUBIC)
 - Low throughput for loss-based CC
 - High packet loss rates when the buffer is smaller than 1.5 bandwidth-delay product (BDP)
 - Throughput variation

1. N. Cardwell et al. "BBR v2, A Model-based Congestion Control." IETF 104, March 2019.

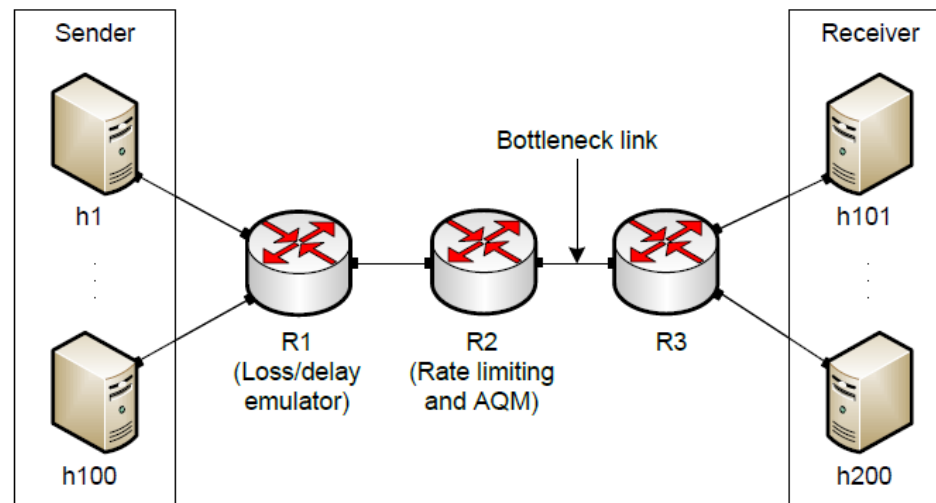
BBRv2 Design Goals

- Improving co-existence with loss-based when sharing a bottleneck link
- Avoiding the bufferbloat problem
- Minimizing the time to reach an equilibrium point (fairness)
- Reducing the variation of the throughput
- Producing high throughput, even with moderate packet losses

1. N. Cardwell et al. "BBR v2, A Model-based Congestion Control." IETF 104, March 2019.

Experimental Setup

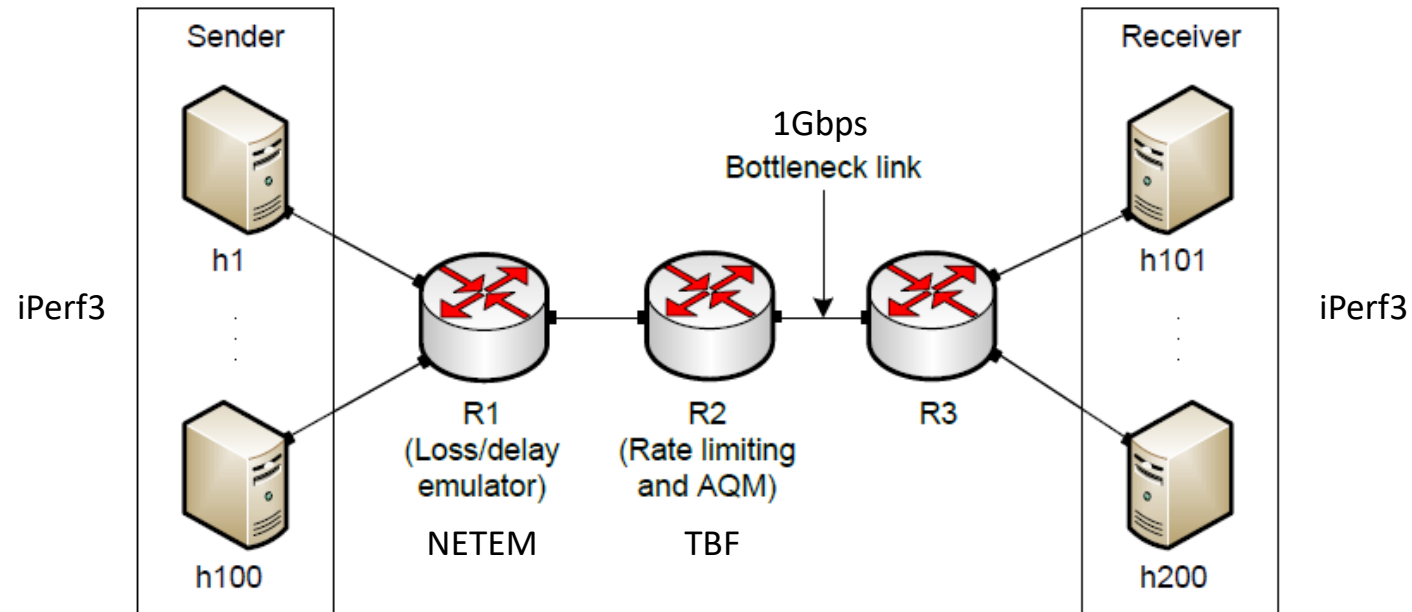
- Up to 100 simultaneous flows
- Mininet network, Linux protocol stack, Tail drop AQM policy (default)
- 300 seconds experiment duration, average results are reported
- CPU cores (8 Xeon 6130) are overprovisioned (below 50% utilization); idle CPU time is a good fidelity indicator, as it indicates that a host is not starved for CPU resources¹



¹N. Handigol et al., "Reproducible Network Experiments Using Container-Based Emulation," ACM CoNEXT '12.

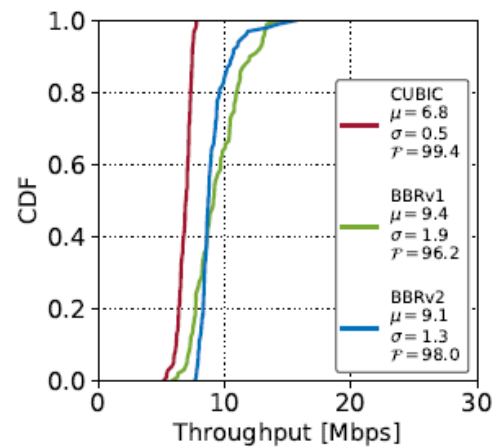
Experimental Setup

- BBRv2 version: v2alpha-2019-07-28
- Bottleneck link = 1Gbps
- All other links = 40Gbps
- Tools: iPerf3, netstat, ss, tc [tbf | netem | fq_codel | cake]

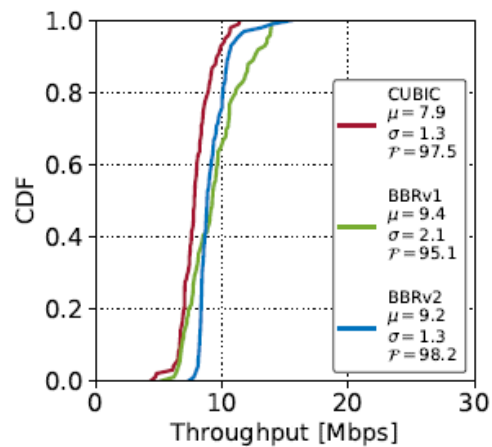


Multiple Flows and Buffer Sizes

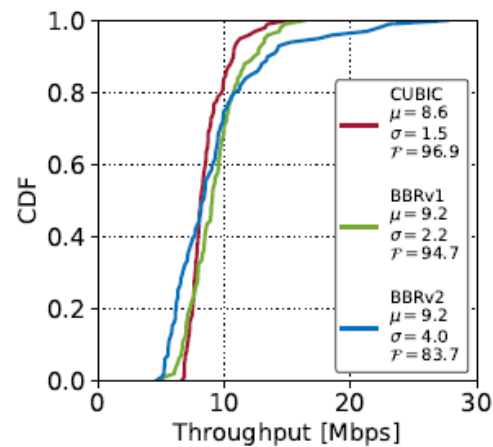
- 100 simultaneous flows, 20ms propagation delay, same CC
- Bottleneck is 1Gbps (ideal allocation is 10Mbps per flow)
- Cumulative distribution function, mean, standard deviation, and fairness



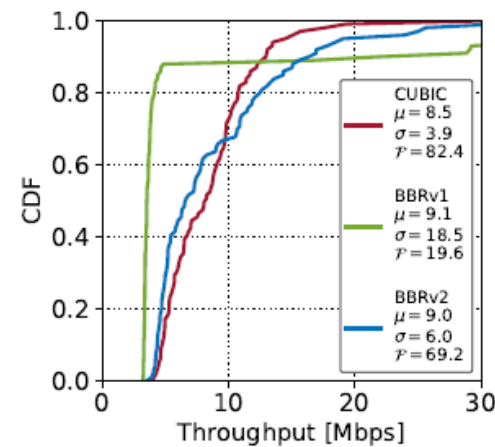
(a) Buffer size: 0.01BDP.



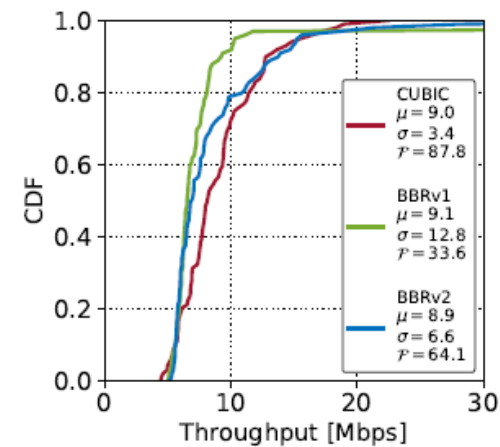
(b) Buffer size: 0.1BDP.



(c) Buffer size: 1BDP.



(d) Buffer size: 10BDP.

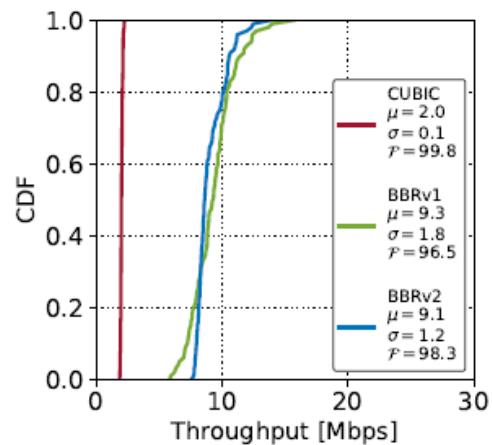


(e) Buffer size: 100BDP.

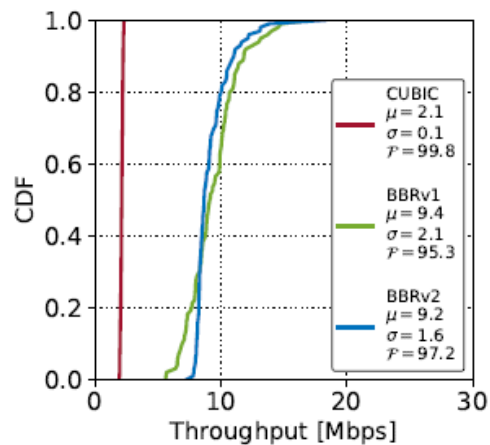
No emulated packet losses

Multiple Flows and Buffer Sizes

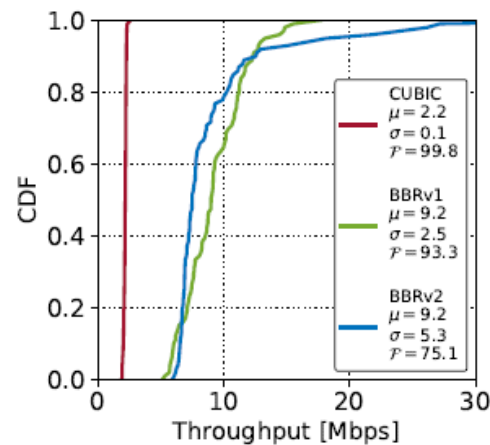
- 100 simultaneous flows, 20ms propagation delay, same CC
- Bottleneck is 1Gbps (ideal allocation is 10Mbps per flow)
- Cumulative distribution function, mean, standard deviation, and fairness



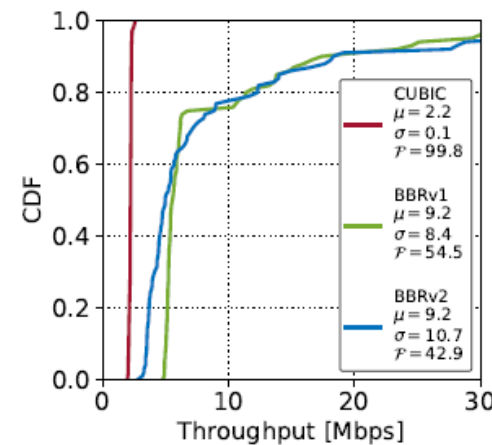
(a) Buffer size: 0.01BDP.



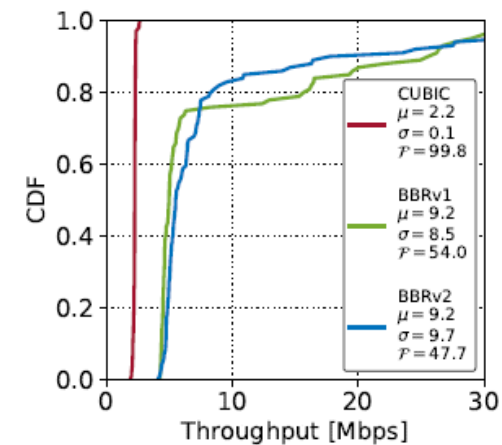
(b) Buffer size: 0.1BDP.



(c) Buffer size: 1BDP.



(d) Buffer size: 10BDP.

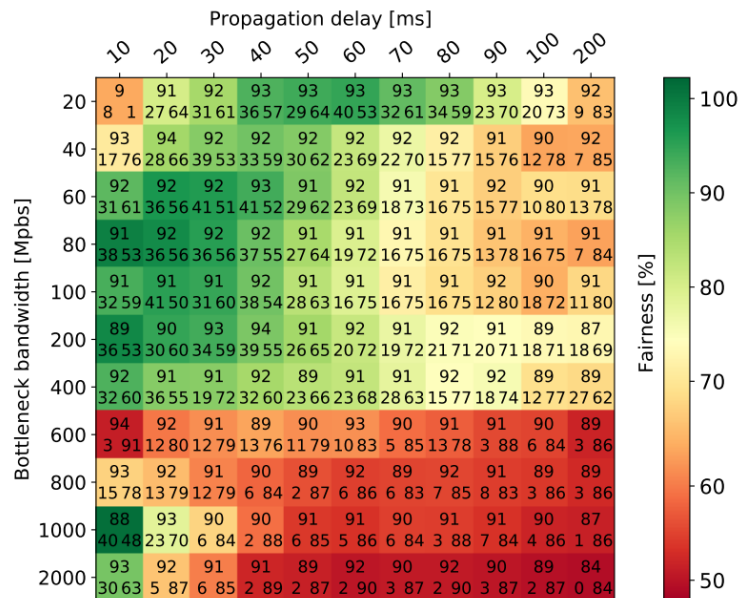


(e) Buffer size: 100BDP.

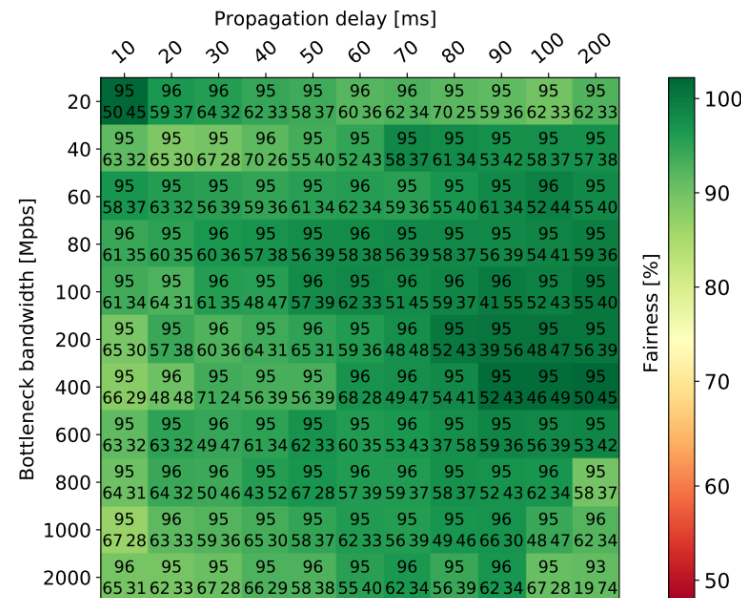
1% packet loss rate

Coexistence and Fairness with CUBIC

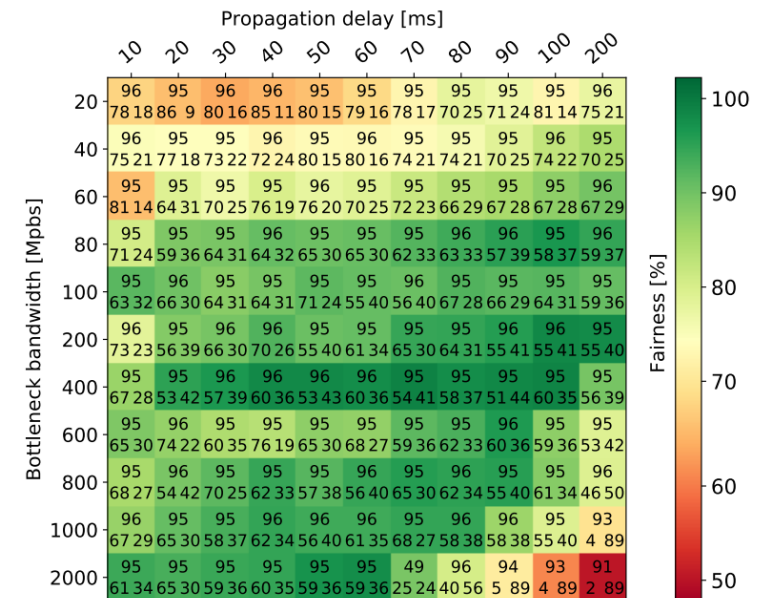
- Fairness index visualized in a heatmap for CUBIC and BBRv2 flows
- 1 CUBIC vs. 1 BBRv2
- Each entry has 3 numbers (in percentage): link utilization (center top), bandwidth used by CUBIC (bottom left value), and bandwidth used by BBRv2 (bottom right)



(a) Buffer size: 0.1BDP.



(b) Buffer size: 1BDP.

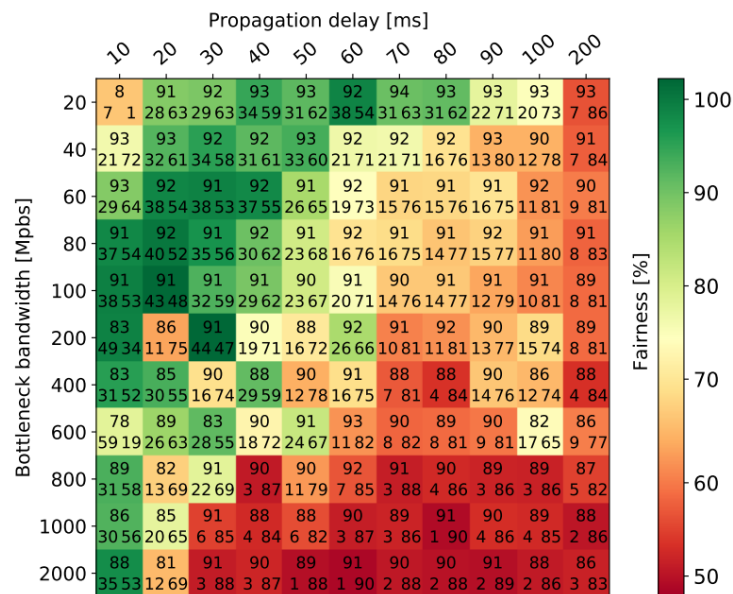


(c) Buffer size: 10BDP.

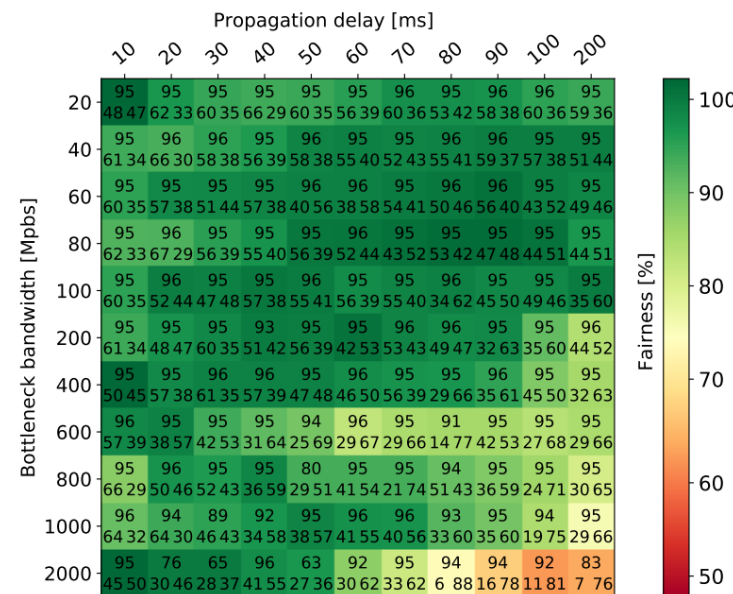
No emulated packet losses

Coexistence and Fairness with CUBIC

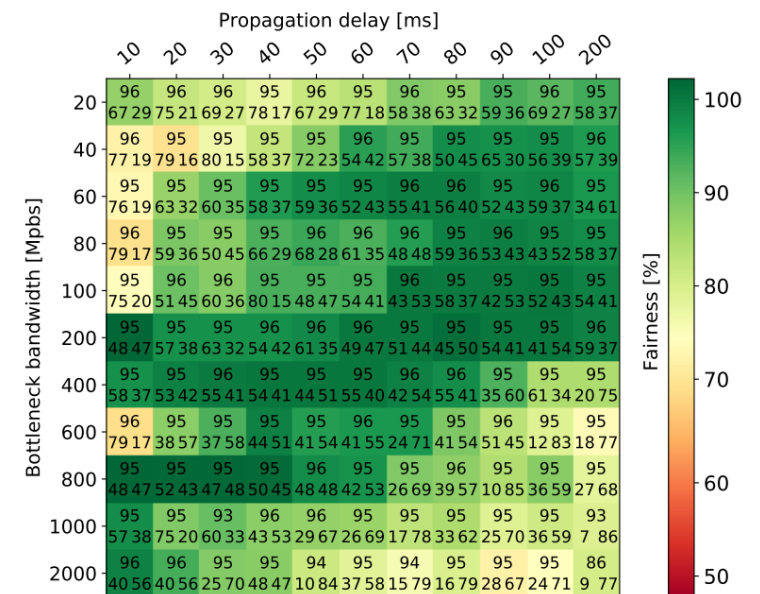
- Fairness index visualized in a heatmap for CUBIC and BBRv2 flows
- 1 CUBIC vs. 1 BBRv2
- Each entry has 3 numbers (in percentage): link utilization (center top), bandwidth used by CUBIC (bottom left value), and bandwidth used by BBRv2 (bottom right)



(d) Buffer size: 0.1BDP.



(e) Buffer size: 1BDP.

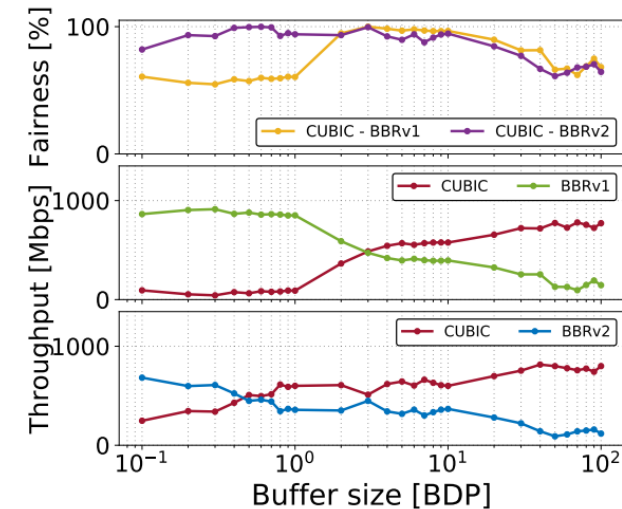


(f) Buffer size: 10BDP.

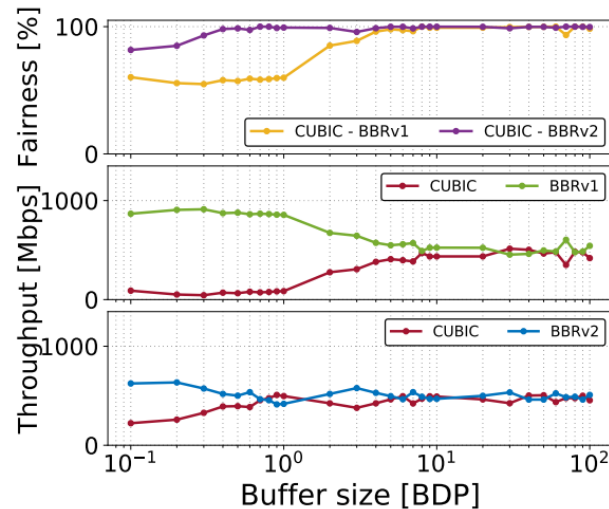
0.01% packet losses

Coexistence and Fairness with CUBIC

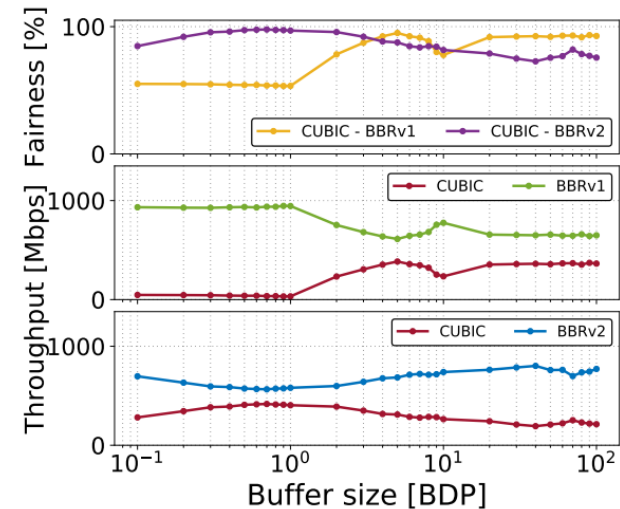
- Bottleneck is 1 Gbps, 20ms propagation delay
- Competing flows sharing the same bottleneck link
- BBRv2 shows a better coexistence with CUBIC than BBRv1



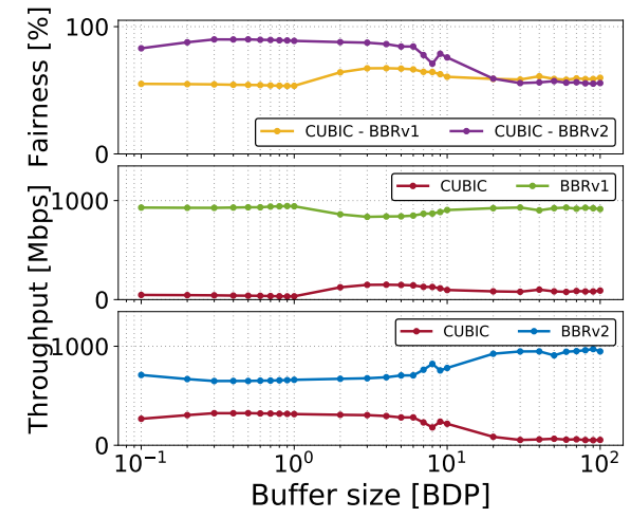
(a) No packet loss, 2 flows.



(b) 0.01% packet loss rate, 2 flows.



(c) No packet loss, 100 flows.

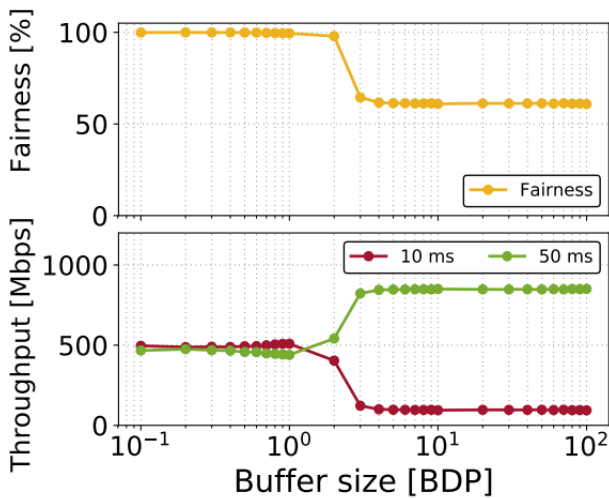


(d) 1% packet loss rate, 100 flows.

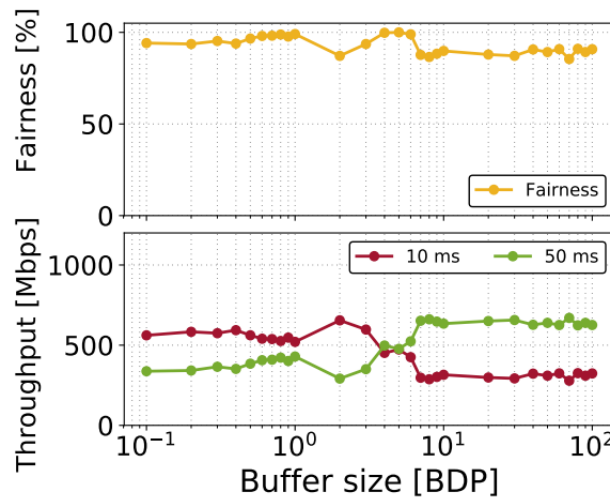
Throughput and fairness as functions of the buffer size

Round-trip Time Unfairness

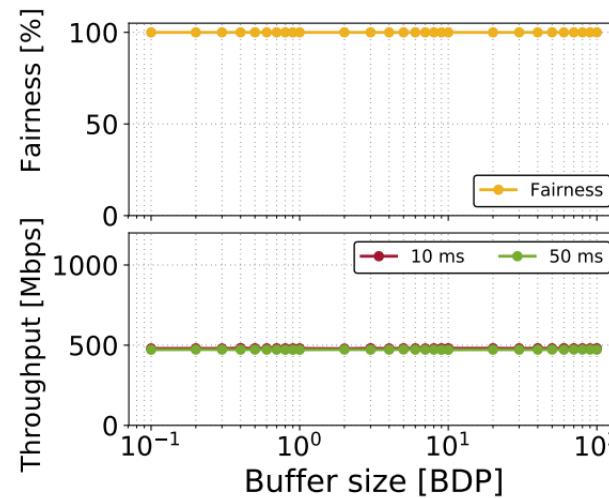
- Two simultaneous flows, one with 10ms propagation delay, and the other with 50ms
- Bottleneck is 1Gbps
- Tail Drop and FQ-CoDel AQMs
- FQ-CoDel mitigates the RTT unfairness



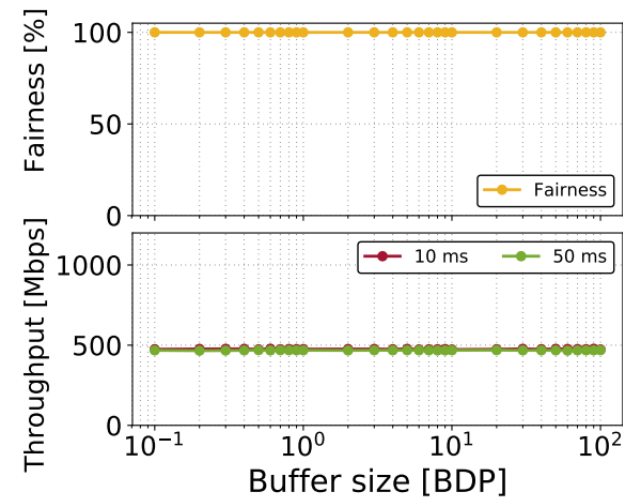
(a) BBRv1 with Tail Drop policy.



(b) BBRv2 with Tail Drop policy.



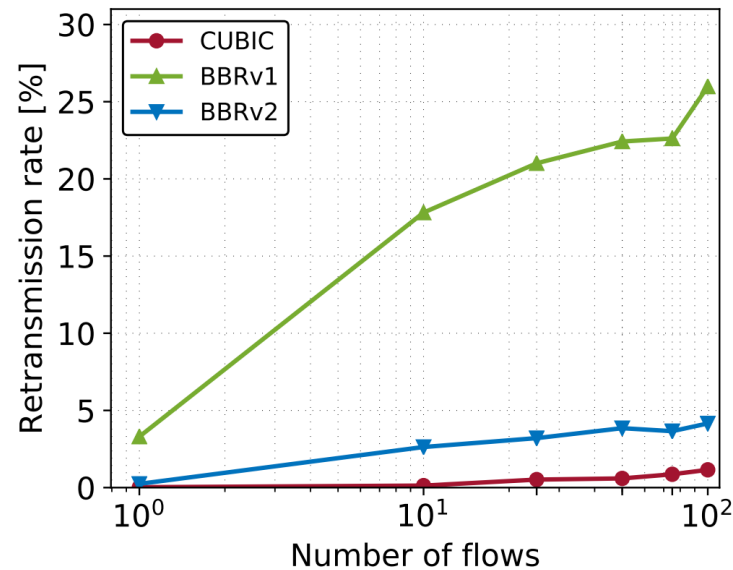
(c) BBRv1 with FQ-CoDel policy.



(d) BBRv2 with FQ-CoDel policy.

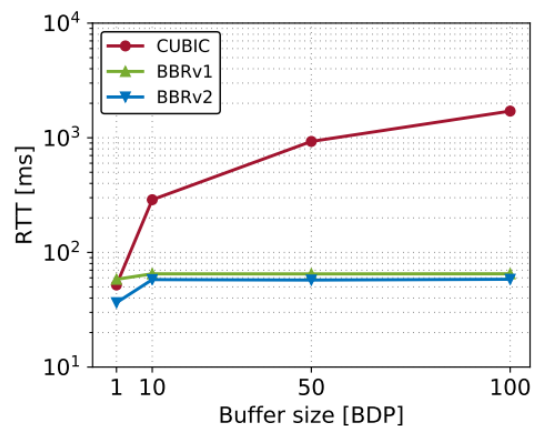
Retransmissions – Number of Flows

- The buffer size is 0.02BDP (i.e., 2.5MB, ~1,666 packets)
- Bottleneck is 1Gbps, 100ms propagation delay
- BBRv1: high retransmission rate with any number of flows (2.5%-25%)
- BBRv2: lower retransmission rate (~2.5%)

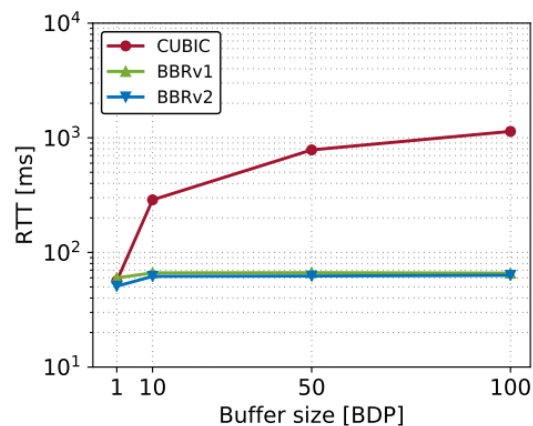


Queueing Delay with Large Buffers

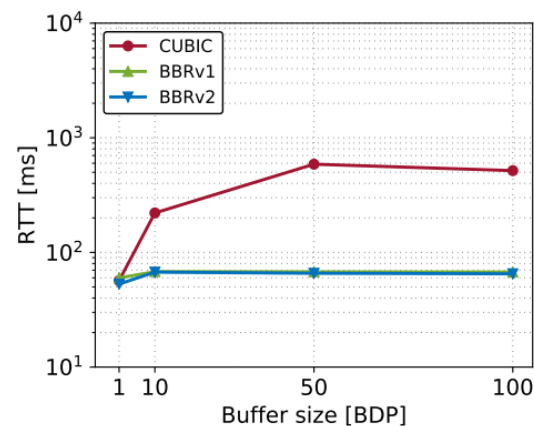
- 2, 10, 25, and 50 simultaneous flows, same CC
- Bottleneck is 1Gbps, 30ms propagation delay
- RTT: queueing delay + propagation delay
- BBRv1 and BBRv2 have low queueing delay independently of the number of flows



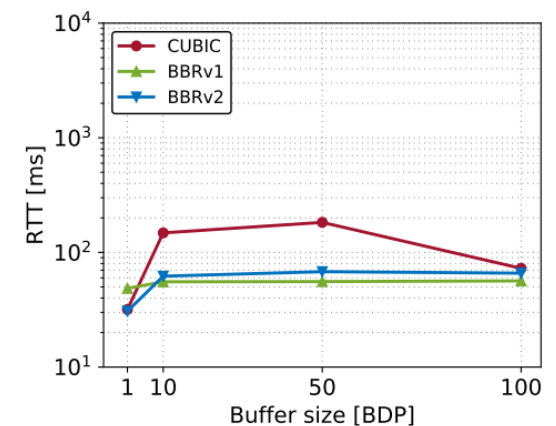
(a) Round-trip time, 2 flows.



(b) Round-trip time, 10 flows.



(c) Round-trip time, 25 flows.

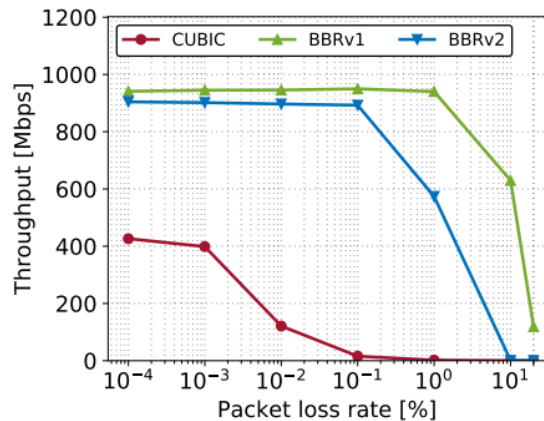


(d) Round-trip time, 50 flows.

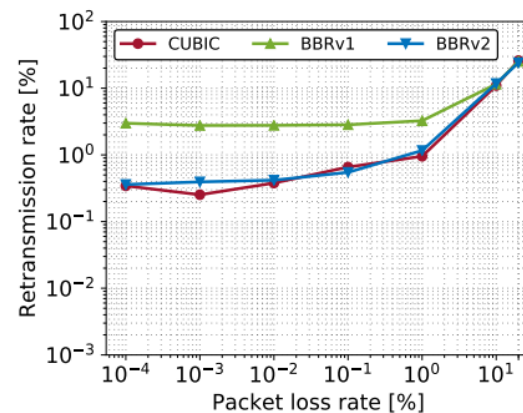
Round-trip time experienced with simultaneous flows

Throughput, Retransmissions – Packet Loss Rate

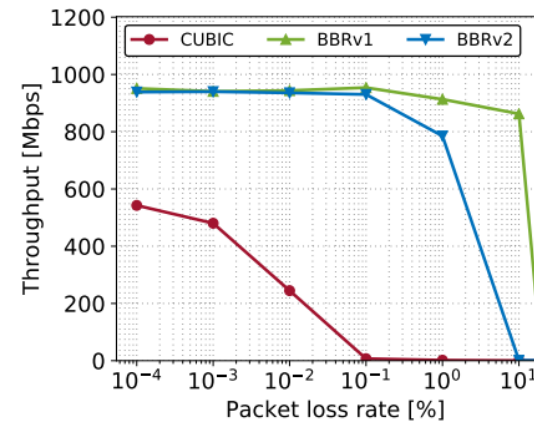
- Single flow
- Bottleneck is 1Gbps, 100ms propagation delay
- BBRv1 is loss-agnostic, which leads to a higher retransmission rate



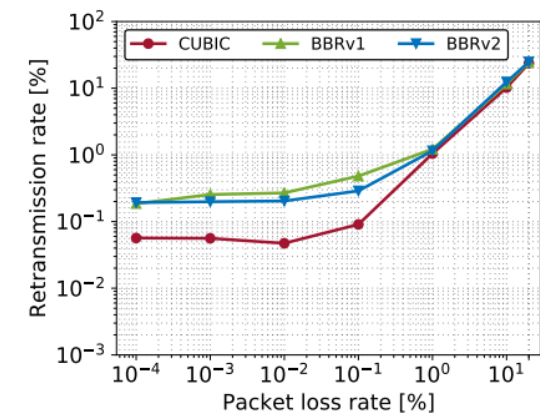
(a) Buffer size: 0.1BDP.



(b) Buffer size: 0.1BDP.



(c) Buffer size: 1BDP.

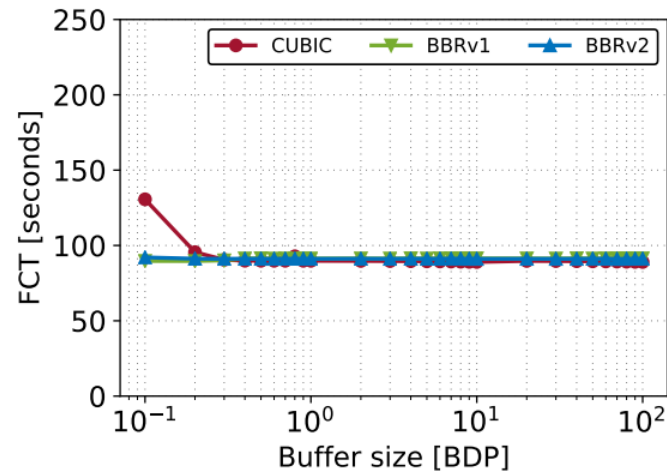


(d) Buffer size: 1BDP.

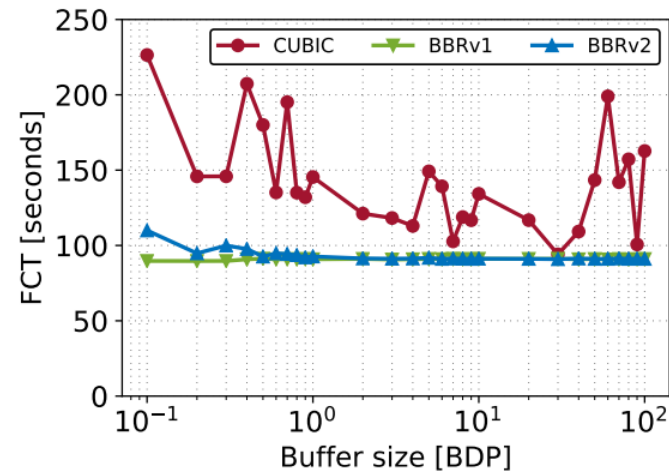
Throughput and retransmission rate as functions of the packet loss rate

Flow-Completion Time (FCT)

- 100 simultaneous flows, same CC
- Bottleneck is 1Gbps, 20ms propagation delay
- Data transfer of 10 GB



(a) No random packet losses.

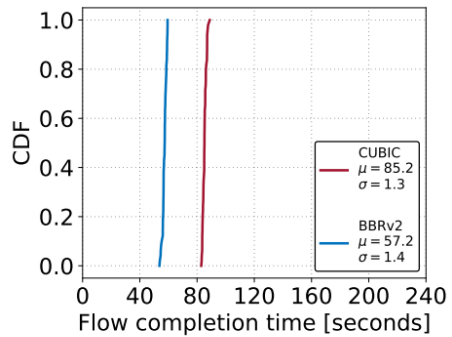


(b) 0.01% random packet loss rate.

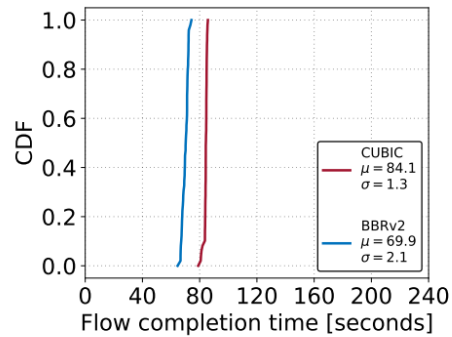
Flow completion time as a function of the buffer size

Flow-Completion Time (FCT)

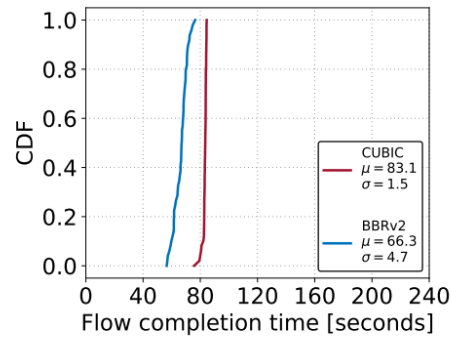
- Bottleneck is 1 Gbps, 20ms propagation delay, 100 simultaneous flows
- 50 CUBIC, 50 BBRv1/BBRv2, each flow completing 100MB of data transfer



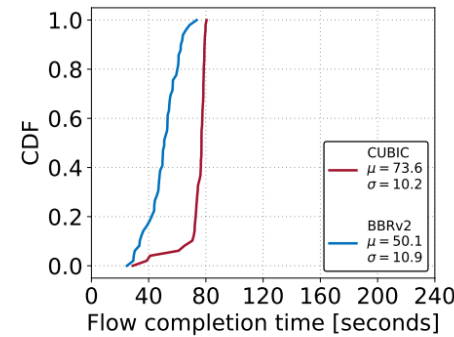
(a) Buffer size: 0.1BDP.



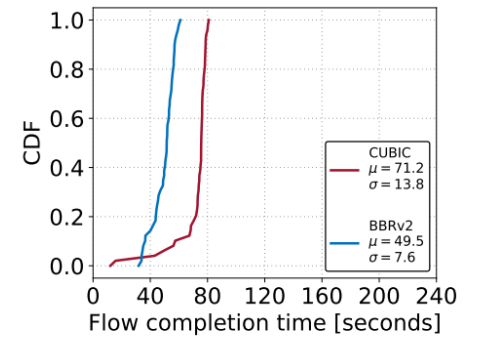
(b) Buffer size: 0.5BDP.



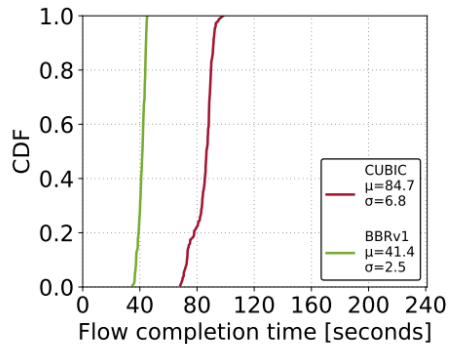
(c) Buffer size: 1BDP.



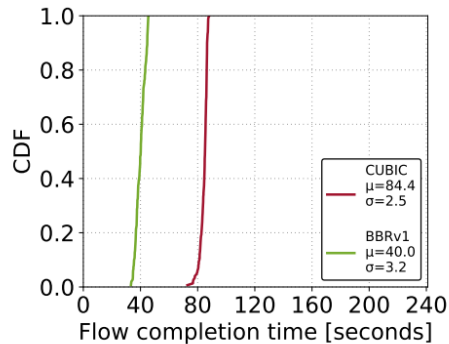
(d) Buffer size: 10BDP.



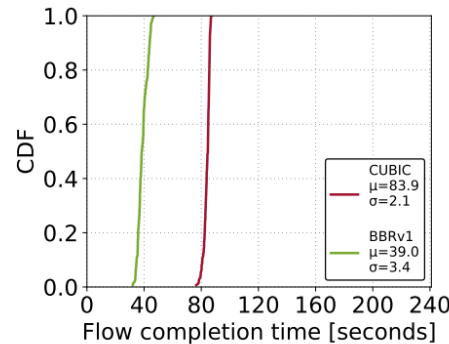
(e) Buffer size: 100BDP.



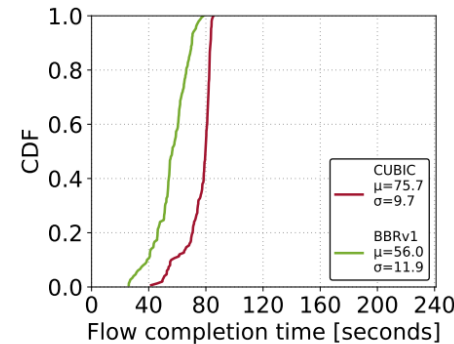
(a) Buffer size: 0.1BDP.



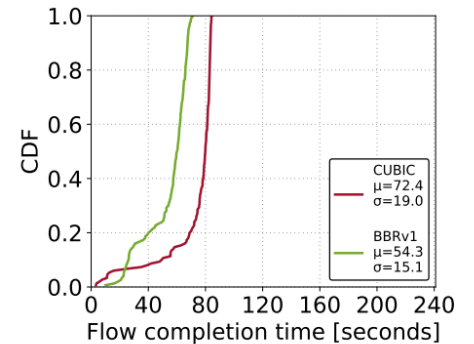
(b) Buffer size: 0.5BDP.



(c) Buffer size: 1BDP.



(d) Buffer size: 10BDP.

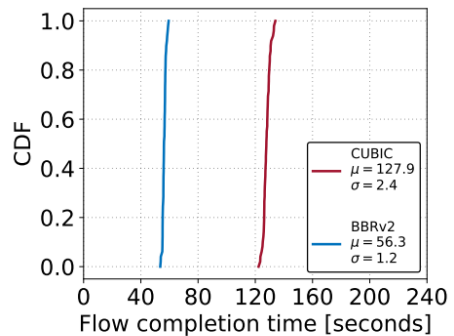


(e) Buffer size: 100BDP.

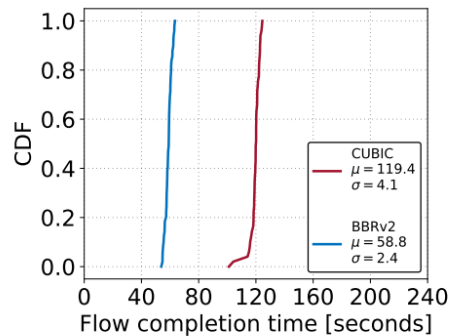
No losses

Flow-Completion Time (FCT)

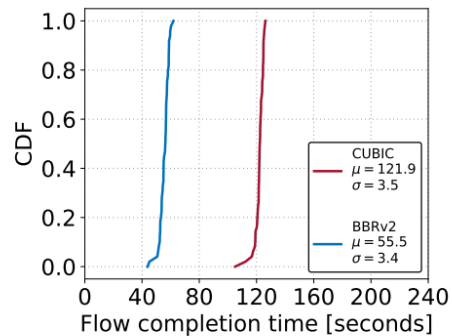
- Bottleneck is 1Gbps, 20ms propagation delay, 100 simultaneous flows
- 50 CUBIC, 50 BBRv1/BBRv2, each flow completing 100MB of data transfer



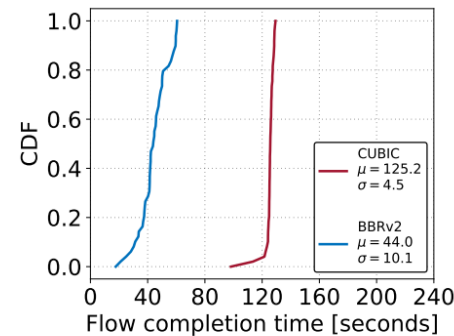
(a) Buffer size: 0.1BDP.



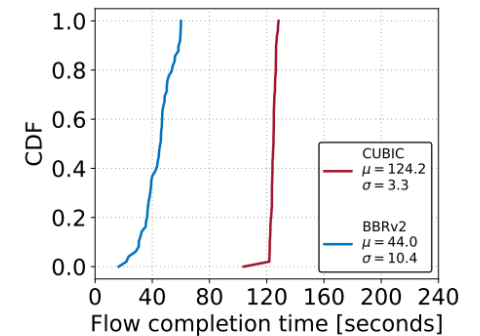
(b) Buffer size: 0.5BDP.



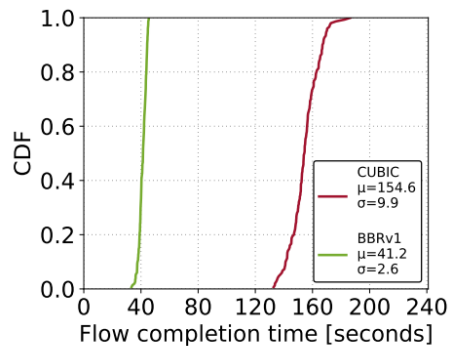
(c) Buffer size: 1BDP.



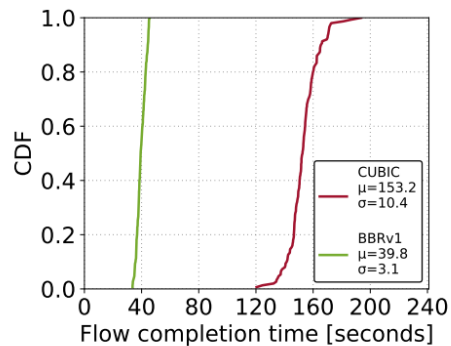
(d) Buffer size: 10BDP.



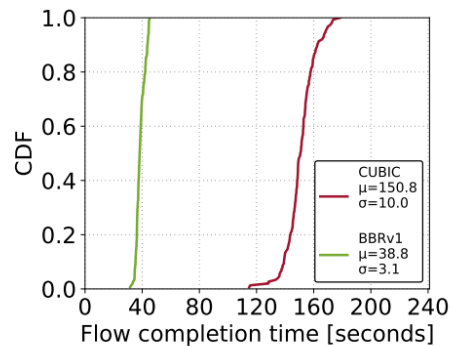
(e) Buffer size: 100BDP.



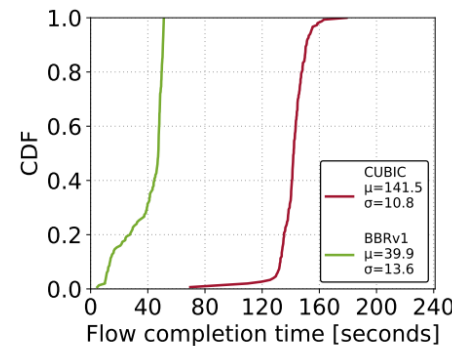
(a) Buffer size: 0.1BDP.



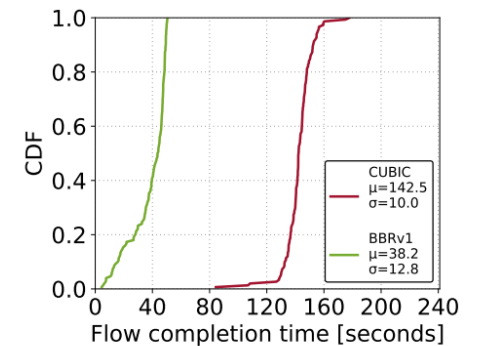
(b) Buffer size: 0.5BDP.



(c) Buffer size: 1BDP.



(d) Buffer size: 10BDP.

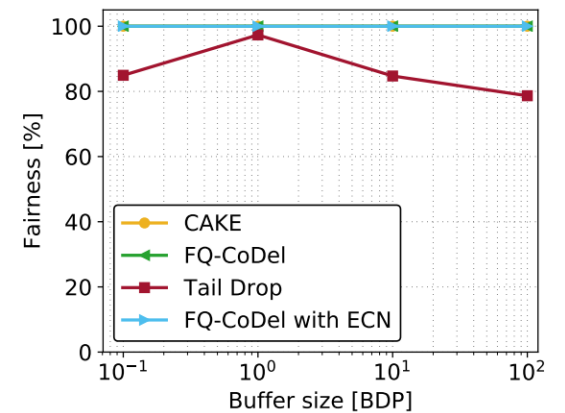
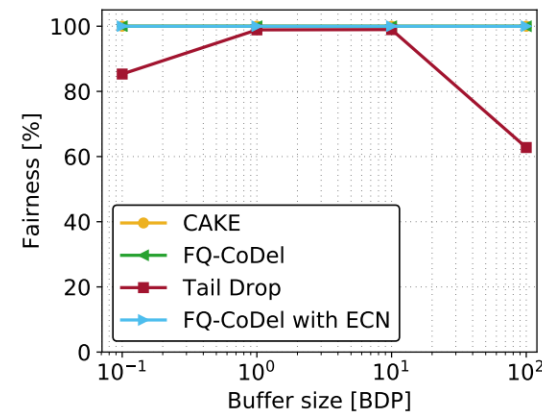
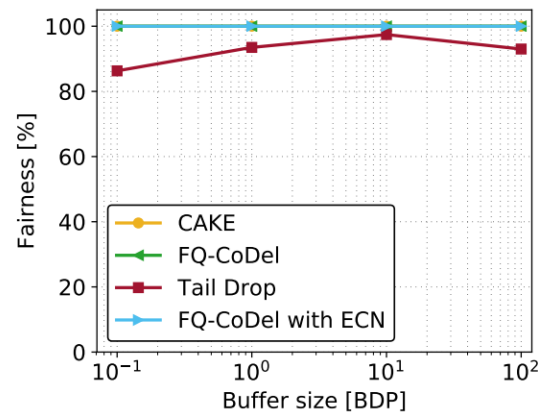
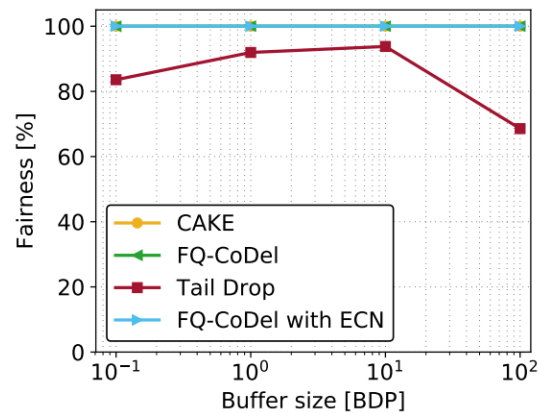


(e) Buffer size: 100BDP.

1% packet losses

Impact of AQM on Fairness

- Bottleneck is 1Gbps, 20ms propagation delay
- 2, 11, 50 and 100 simultaneous flows, uneven composition
- Tail drop, CAKE, FQ_CoDel, FQ_CoDel w/ ECN



(a) 1 BBRv2 flow, 1 CUBIC flow.

(b) 1 BBRv2 flow, 10 CUBIC flows.

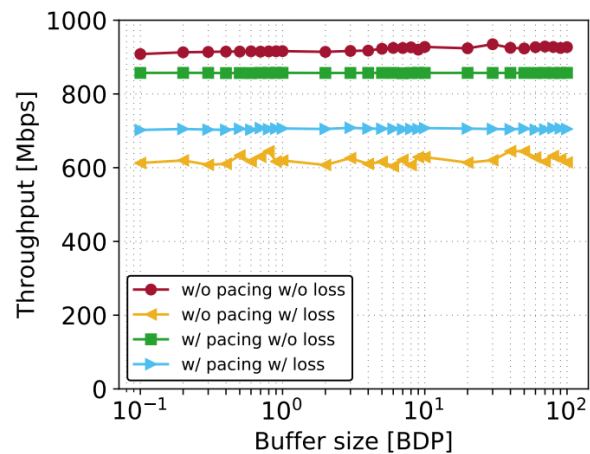
(c) 10 BBRv2 flows, 10 CUBIC flows.

(d) 50 BBRv2 flows, 50 CUBIC flows.

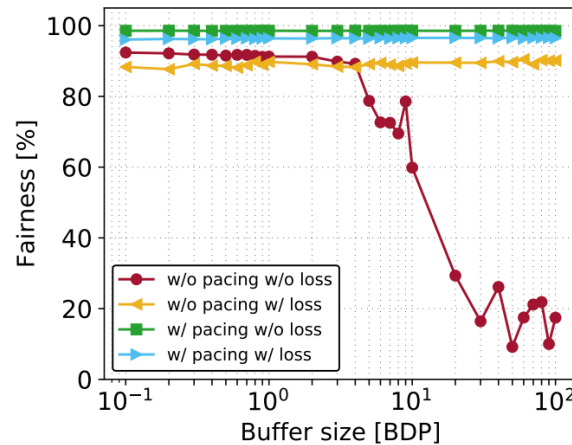
No losses

The Effects of Fixed-rate Pacing on TCP CUBIC

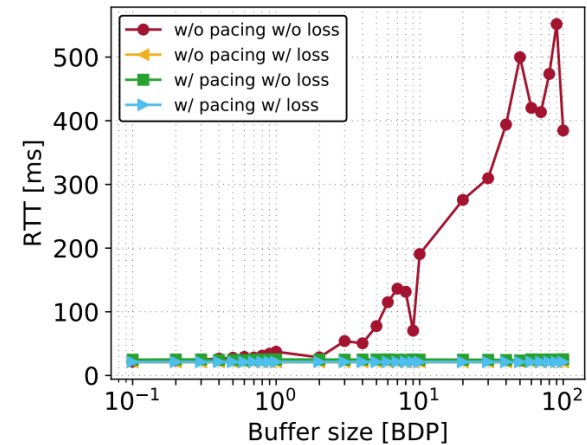
- 100 simultaneous flows
- Bottleneck is 1Gbps, 20ms propagation delay
- Pacing rate per flow is 8.5Mbps, 85% link utilization



(a) Throughput.



(b) Fairness index.



(c) Round-trip time.

No losses and with 1% packet losses

Conclusions

- BBRv2 addressed the limitations of BBRv1
 - Tolerates much higher packet loss rates than CUBIC, lower than BBRv1
 - Mitigates the unfairness problem
 - Has better coexistence with CUBIC than BBRv1
 - Produces lower retransmission rates than BBRv1
 - Exhibits low queueing delay, even with bloated buffers
 - Coexistence / fairness problem with CUBIC is not solved but mitigated

Resources

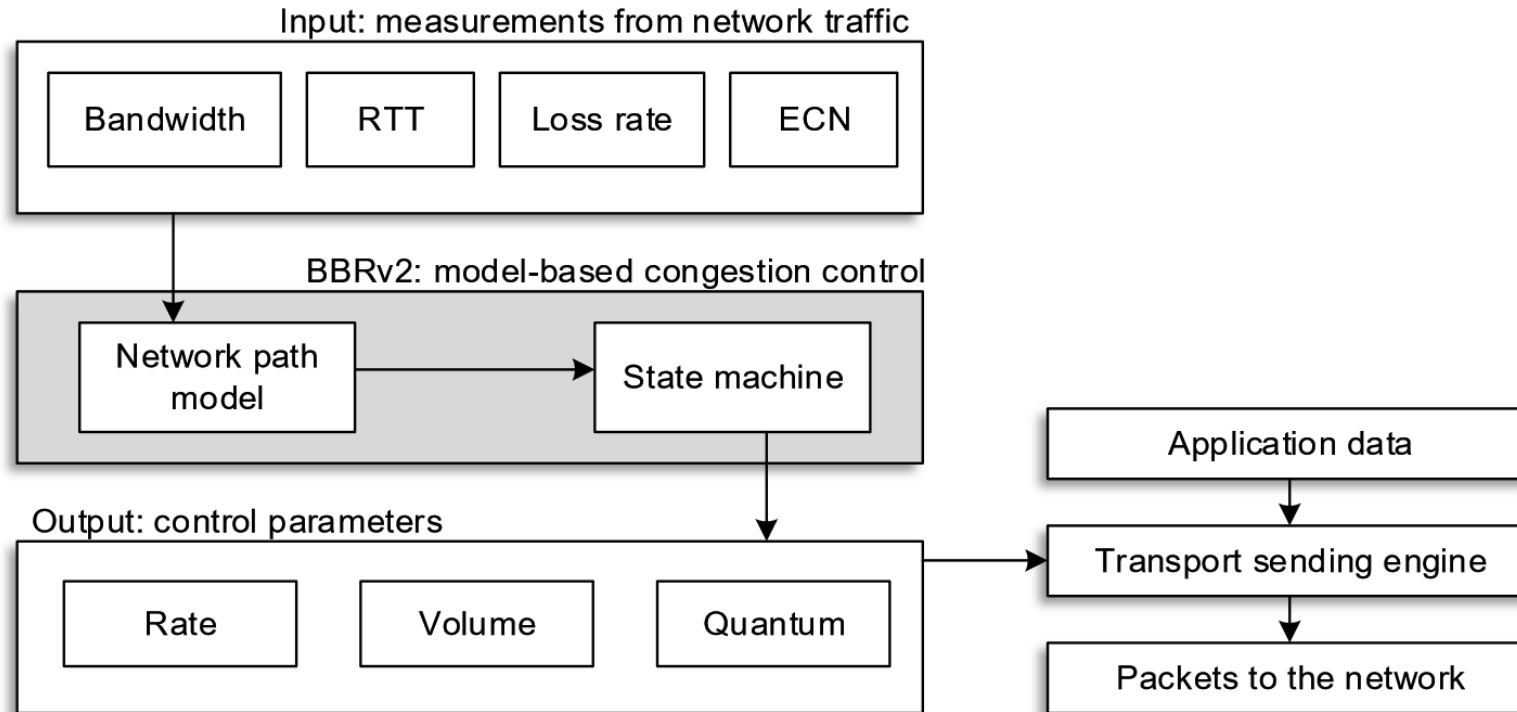
- Full paper:
<https://www.sciencedirect.com/science/article/pii/S014036642030092X>
- The scripts for emulation are available in the following GitHub repository
<https://github.com/gomezgaona/bbr2>
- BBRv2 version: v2alpha-2019-07-28
- Cyberinfrastructure website:
<http://ce.sc.edu/cyberinfra/>
- BBRv2 GitHub repository:
<https://github.com/google/bbr/tree/v2alpha>



UNIVERSITY OF
SOUTH CAROLINA

BBRv2

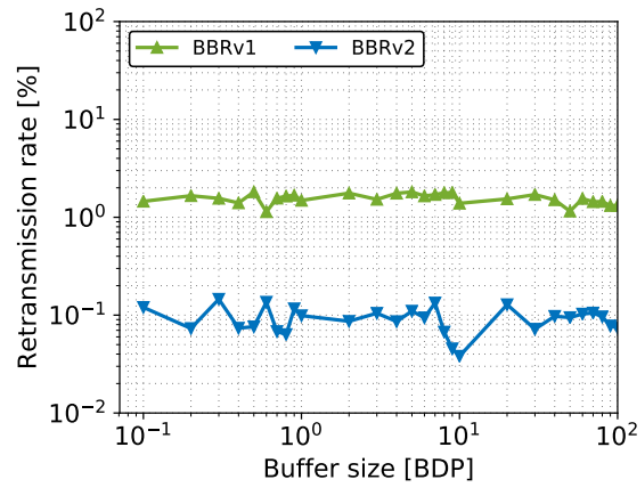
- BBRv2 architectural diagram¹



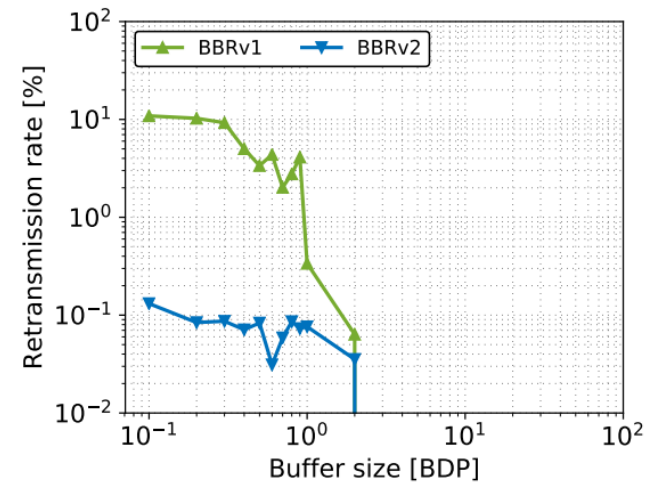
1. N. Cardwell et al. "BBR v2, A Model-based Congestion Control." IETF 104, March 2019.

Retransmission Rate

- Two simultaneous flows, one with 10ms prop delay, and the other with 50ms
- Bottleneck is 1Gbps
- BBRv2 produces a low retransmission rate, (e.g., ~0.1%)



(a) Retransmissions, Tail Drop.

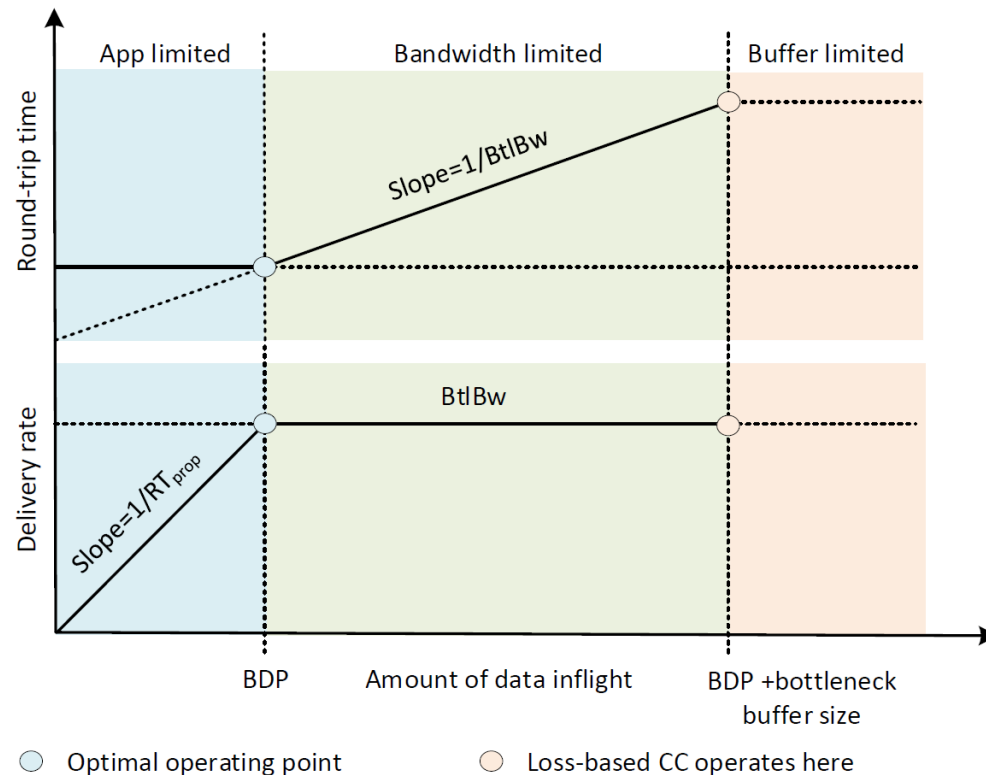


(b) Retransmissions, FQ-CoDel.

Retransmission rate generated by BBRv1 and BBRv2

BBR- Model-based Congestion Control

- BBR: **B**ottleneck **B**andwidth and **R**ound-trip propagation time¹
- Dynamically estimates windowed max BW and min RTT on each ACK



1. N. Cardwell et al. "BBR: Congestion-Based Congestion Control: Measuring bottleneck bandwidth and round-trip propagation time." Queue 14.5 (2016): 20-53..