IETF 116
ICCRG

29th March 2023

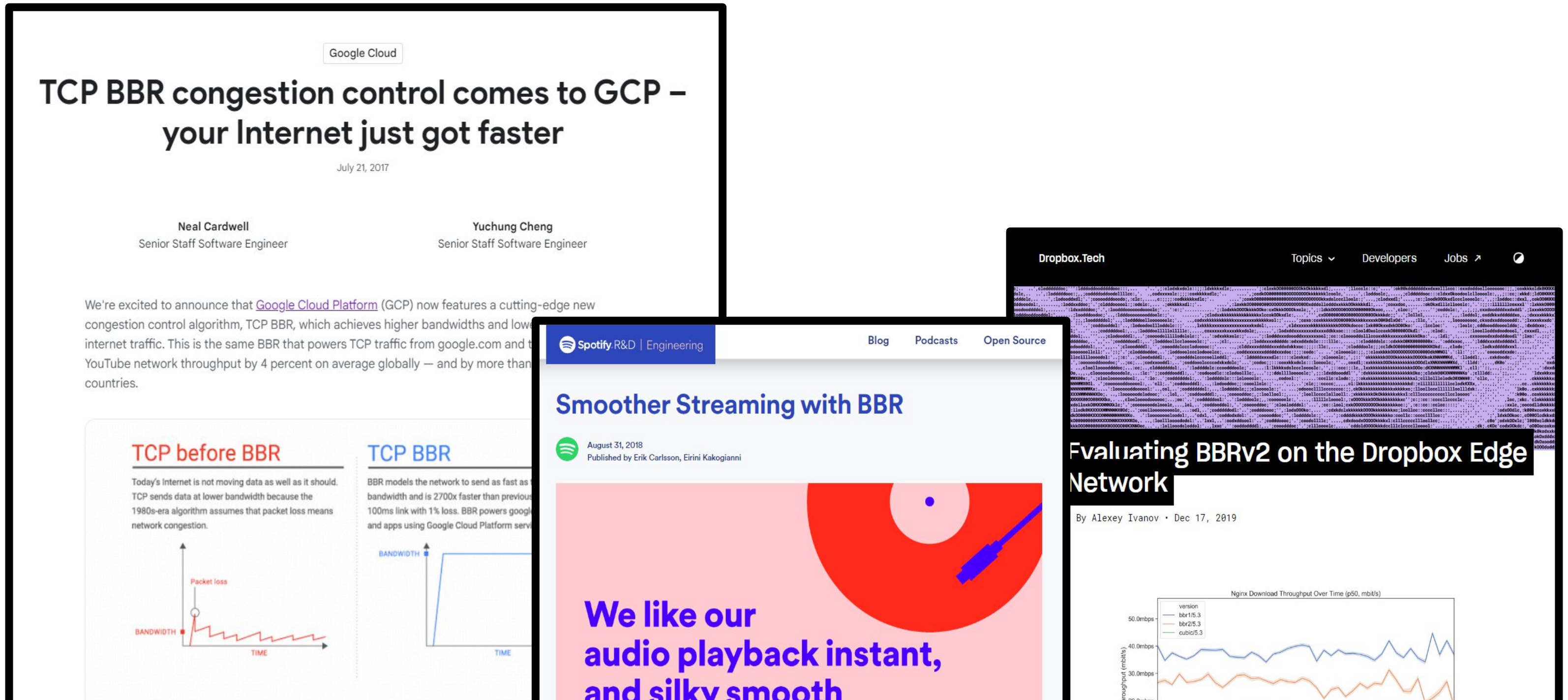
Are we heading towards a BBR-dominant Internet?

Ayush Mishra, Tiu Wee Han, Ben Leong

National University of Singapore

The promise of higher bandwidths!

Google, Dropbox, and Spotify are reporting **higher throughput** and **lower delay jitter** after switching to BBR



BBR's Rapid adoption

Variant	Websites	Proportion
CUBIC [15]	6,139	30.70%
BBR [4]	3,550	17.75%
BBR G1.1	167	0.84%
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]

(Mishra et al, SIGMETRICS 2020)

→ In three short years, BBR already accounted for **18%** of the top 20,000 websites on the Internet.

Traffic share estimated around **40%**

If you run a website and care about **throughput**, it is natural to consider switching from CUBIC to BBR.

Are we heading towards an **all-BBR Internet** then?

Are we heading towards an **all-BBR Internet** then?

#1 How will BBR's throughput gains over CUBIC evolve as more people switch to BBR?

Mathematical model

Are we heading towards an **all-BBR Internet** then?

#1 How ... R's
thr ...
over ...
more pe ... tch
to BBR?

Game Theory

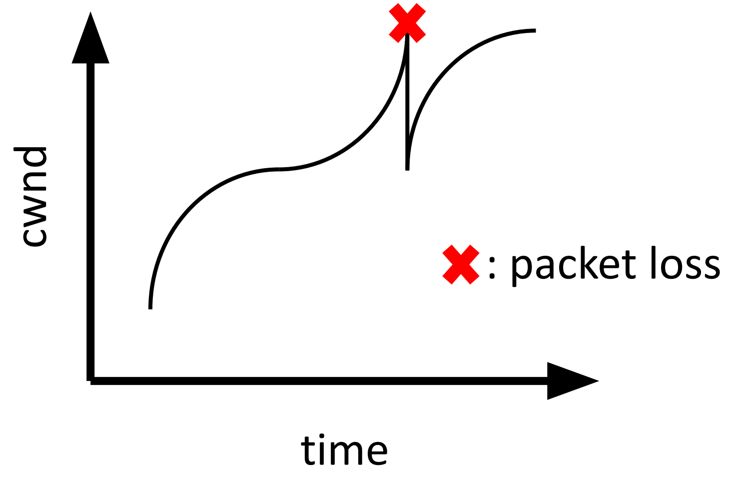
#2 How will these evolving throughput gains dictate the future CCA landscape?

A primer on CUBIC

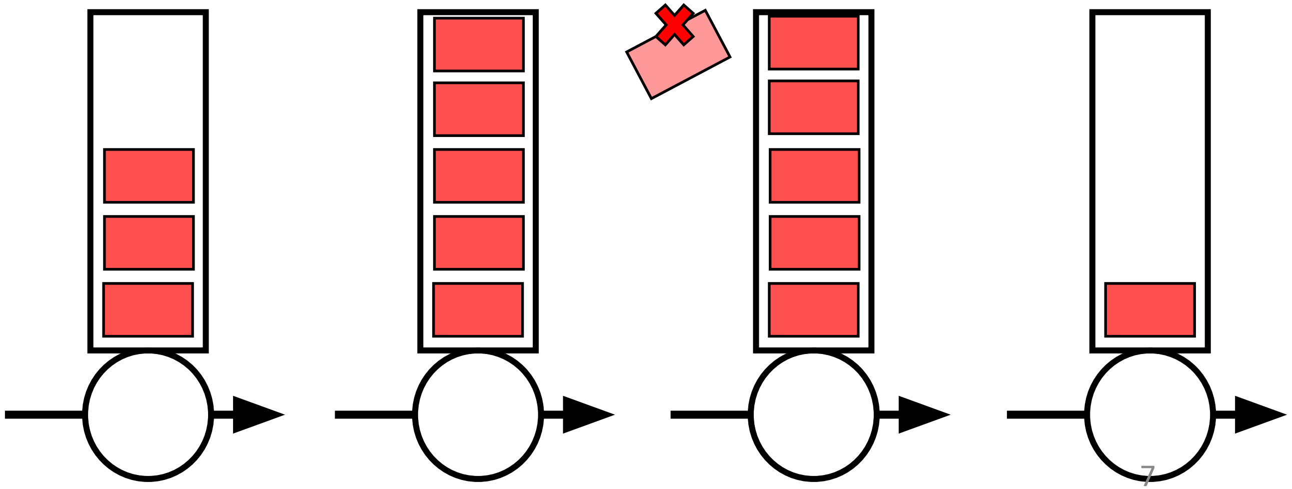
Cwnd-based congestion control algorithm

Treats **packet loss** as a congestion signal.

Reduces cwnd by 30% when is sees a packet loss.



Considered a **buffer filler**

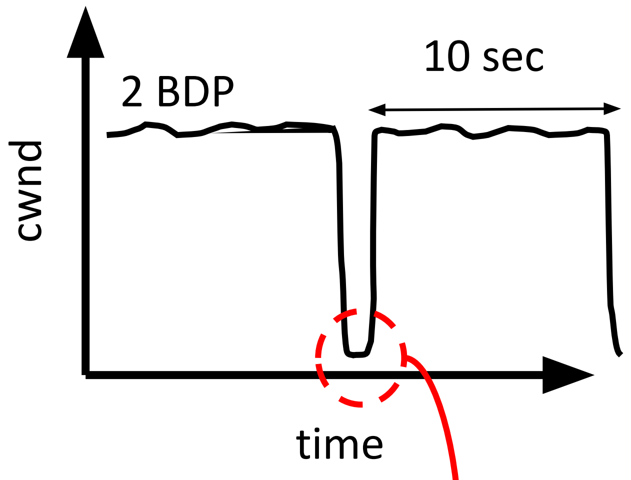


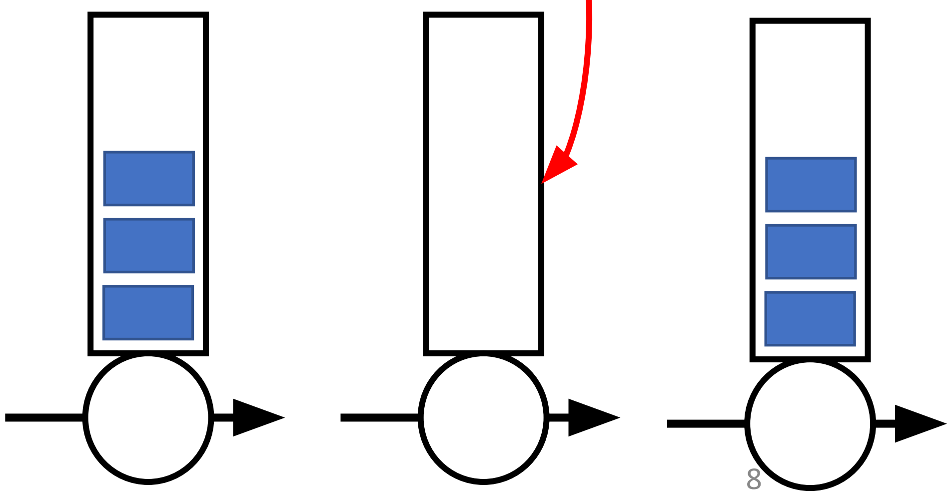
A primer on BBR

Rate-based congestion control algorithm. Uses **RTT$_{min}$** and **bandwidth** estimates to infer congestion.

Becomes **cwnd-limited** when it competes with CUBIC*. **cwnd = 2 BDP**

Backs off every 10 sec to measure RTT$_{min}$

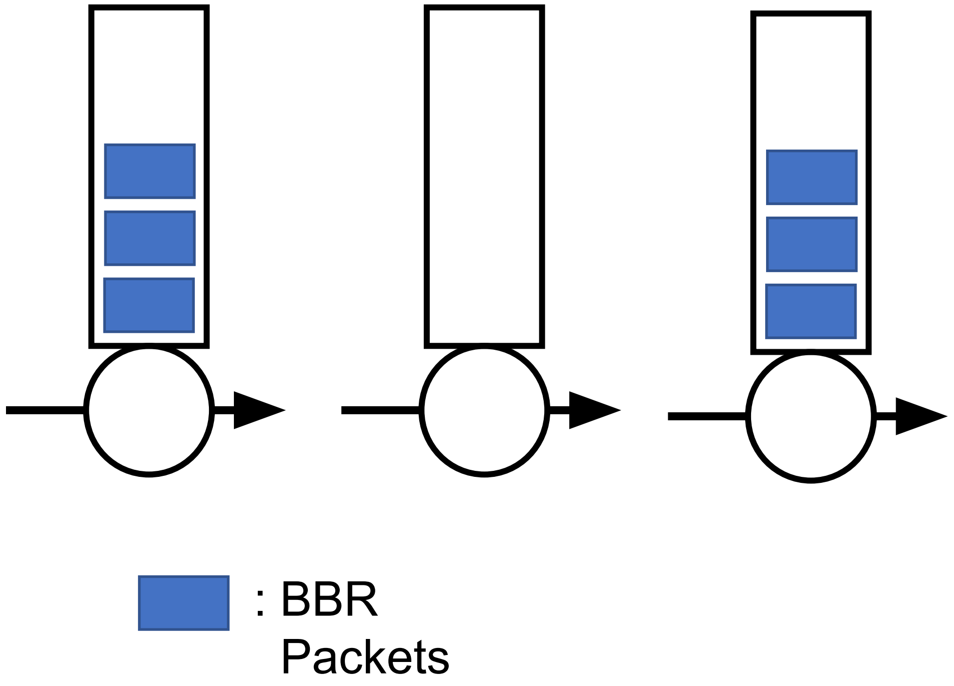




*according to Ware et al, IMC 2019

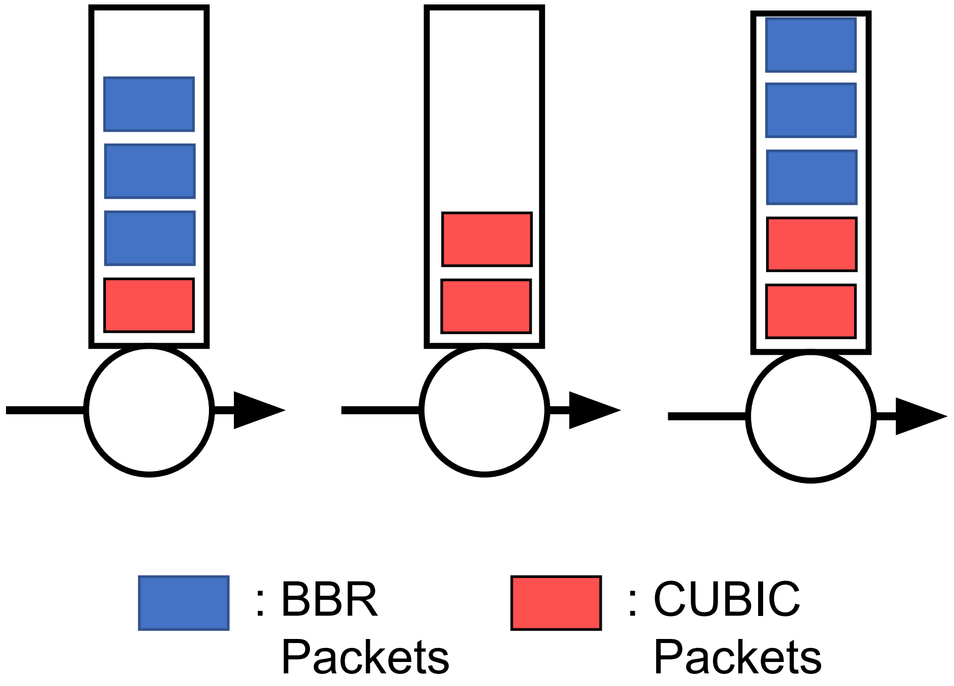
RTT_{min} overestimation

BBR wants to empty the buffer every 10 sec



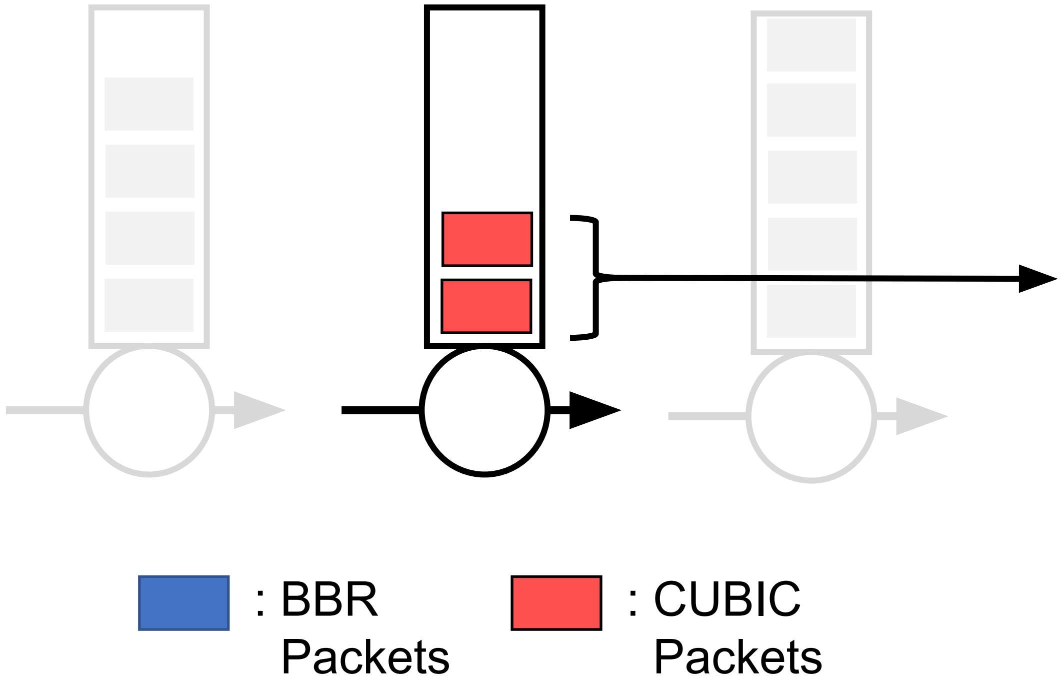
RTT_{min} overestimation

But BBR can't empty the buffer every 10 seconds because of CUBIC's packets!



RTT_{min} overestimation

But BBR can’t empty the buffer every 10 seconds because of CUBIC’s packets!



This leads to **RTT_{min} overestimation** for BBR

Basic 2-flow model

5 key assumptions

1. All competing flows have the **same RTT**
2. The buffer is at least 1 BDP and the **link is always utilized**
3. BBR always has **2 BDP packets in flight**
4. Packets are **uniformly distributed** and the buffer is **droptail**
5. BBR’s reduction in bandwidth while probing for RTT_{min} is negligible

Basic 2-flow model

1

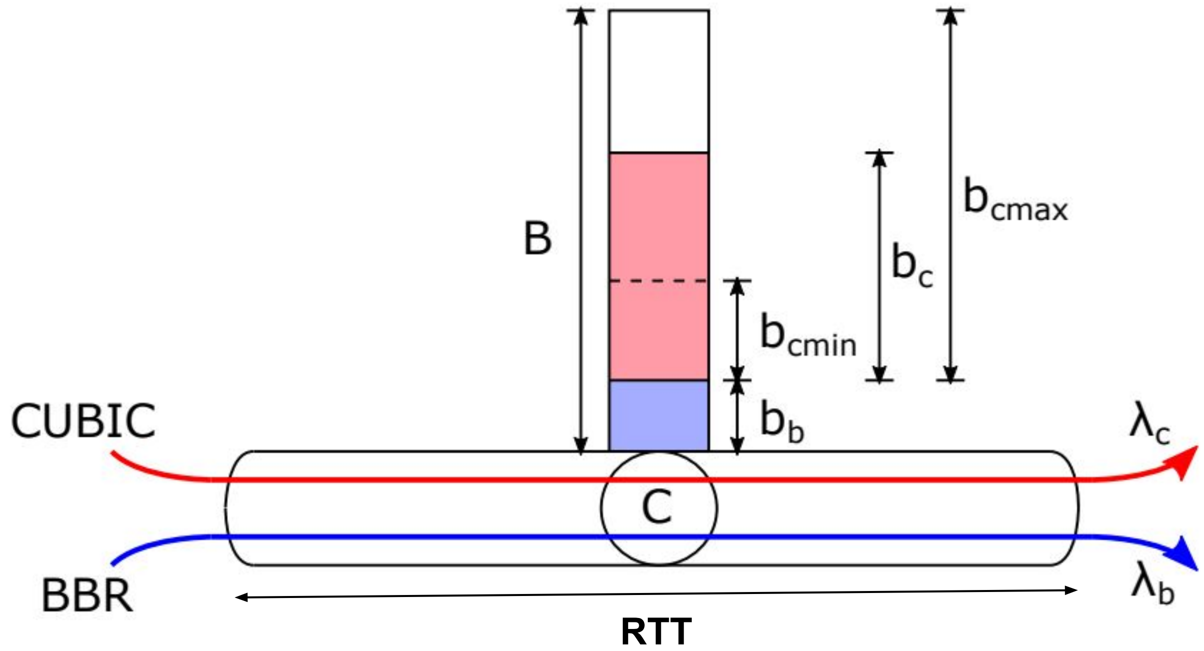
BBR's throughput is cwnd divided by delay

$$\lambda_b \leftarrow \frac{2\lambda_b RTT^+}{RTT + Q_d}$$

2

Where RTT⁺ is BBR's over estimated RTT because of CUBIC

$$RTT^+ = RTT + \frac{b_{cmin}}{C}$$



3

Extent of this overestimation is based on CUBIC's back off:

$$b_{cmin} = (0.7W_{max}) - (\lambda_{cmin} RTT)$$

Basic 2-flow model

1

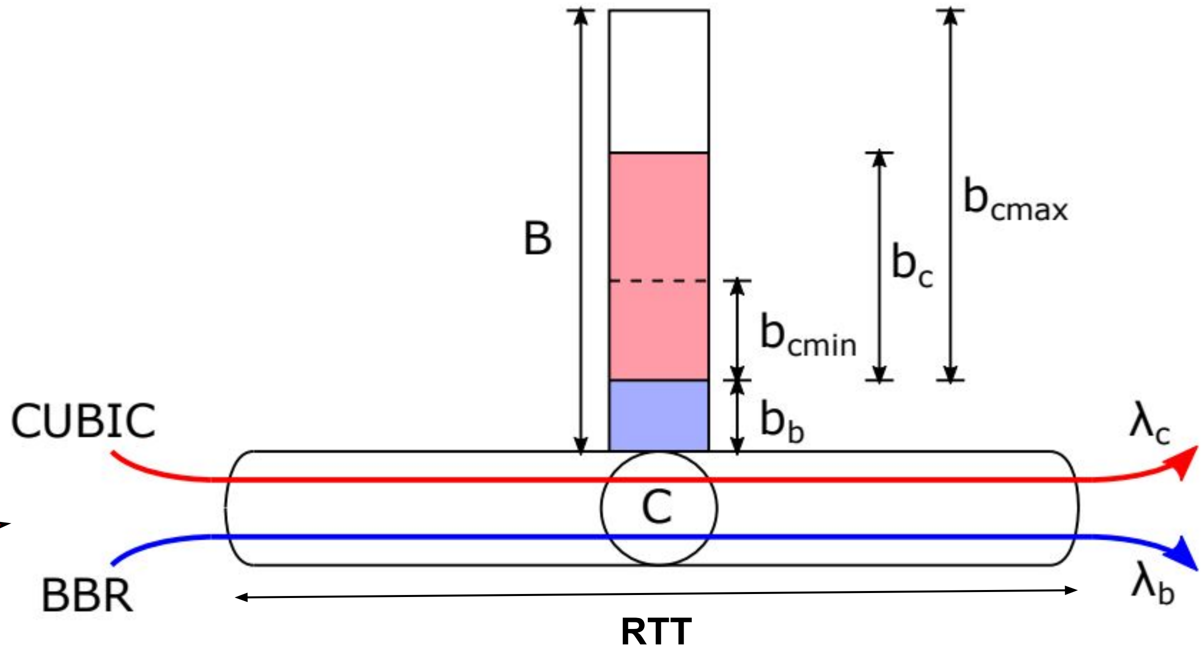
BBR's throughput is cwnd divided by delay

$$\lambda_b \leftarrow \frac{2\lambda_b RTT^+}{RTT + \ldots}$$

Solve!

be… CUBIC

$$RTT^+ = RTT + \frac{b_{cmin}}{C}$$



3

Extent of this overestimation is based on CUBIC's back off:

$$b_{cmin} = (0.7W_{max}) - (\lambda_{cmin}RTT)$$

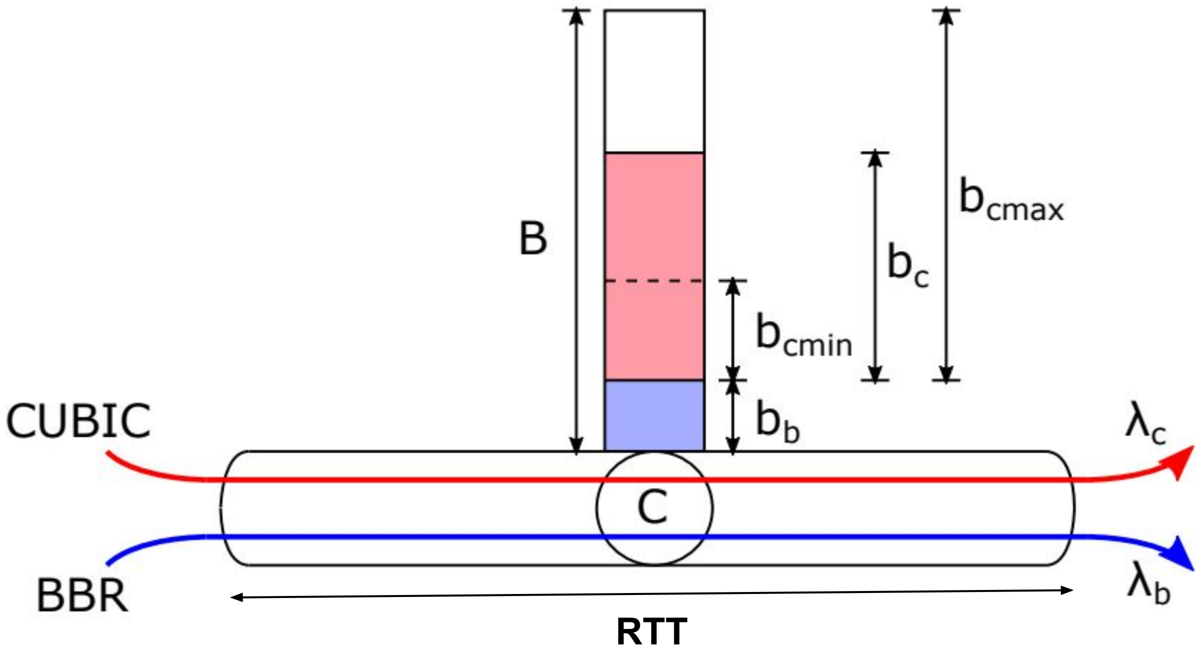
Basic 2-flow model

CUBIC's throughput

$$\lambda_c = \frac{2b_{cmin} + C \cdot RTT - B_b}{\left(RTT + \frac{2b_{cmin}}{C}\right)}$$

BBR's throughput

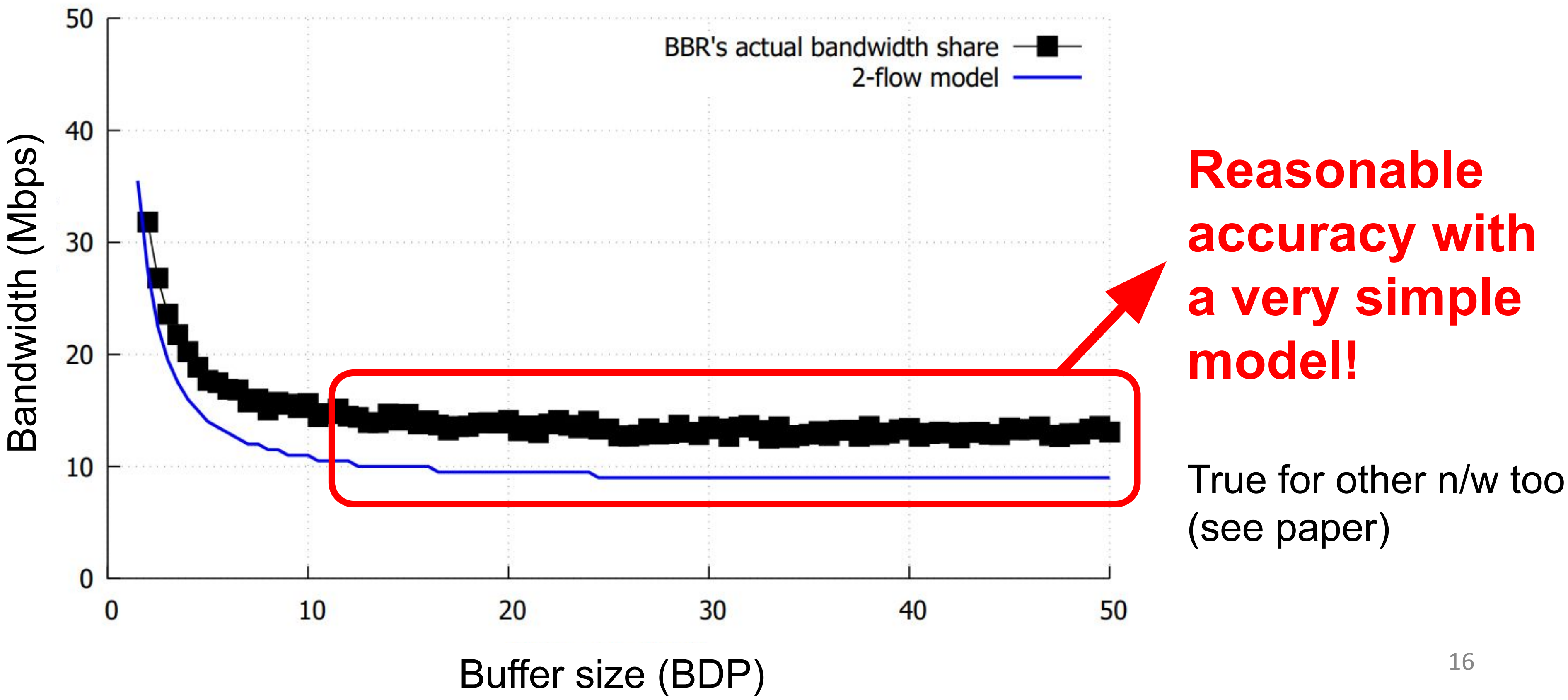
$$\lambda_b = C - \lambda_c$$



Validating the 2-flow model

Ran a CUBIC and a BBR flow through a 50 Mbps link with 40 ms RTT

Plotted the empirical and predicted throughput across buffer sizes



True for other n/w too (see paper)

Extending the model to multiple flows

Basic 2-flow model:

Extent of this overestimation is based on CUBIC's back off behavior

$$b_{cmin} = (0.7W_{max}) - (\lambda_{cmin}RTT)$$

No longer true for multiple CUBIC flows!

RTT overestimation now also depends on the degree of synchronization between the CUBIC flows.

Extending the model to multiple flows

Basic 2-flow model:

Extent of this overestimation is based on CUBIC's back off behavior

$$b_{cmin} = (0.7W_{max}) - (\lambda_{cmin}RTT)$$

Solve again!

Solution:
Predict the upper and lower bounds instead

Sync bound

$$\hat{b}_{cmin} = (0.7\hat{W}_{max}) - (\hat{\lambda}_{cmin}RTT)$$

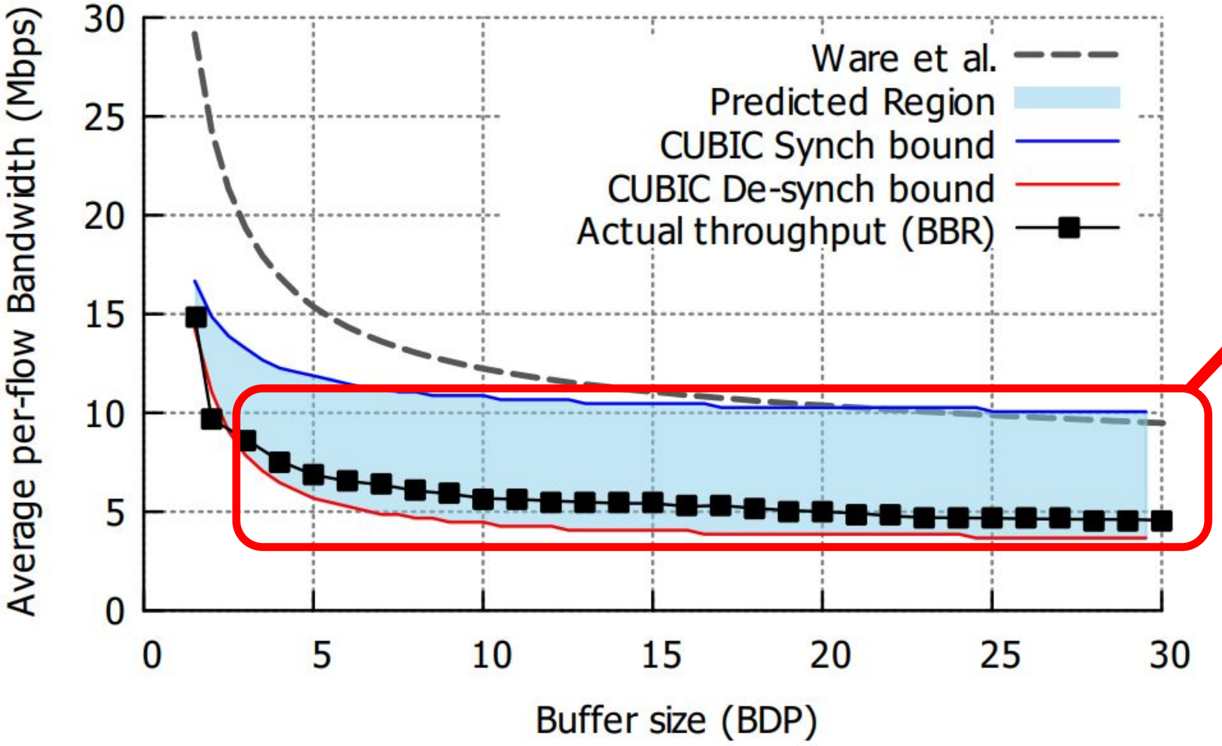
De-sync bound

$$\hat{b}_{cmin} = \left(\frac{(N_c - 0.3)}{N_c}\hat{W}_{max}\right) - (\hat{\lambda}_{cmin}RTT)$$

Validating the multiple flow model

Launched 5 CUBIC and 5 BBR flows through a 100 Mbps 40 ms link

Plotted the empirical and predicted throughput across buffer sizes



Actual throughput is within predicted bounds

True for other n/w too
(see paper)

Are we heading towards an **all-BBR Internet** then?

#1 How will BBR's throughput gains over CUBIC evolve as more people switch to BBR?

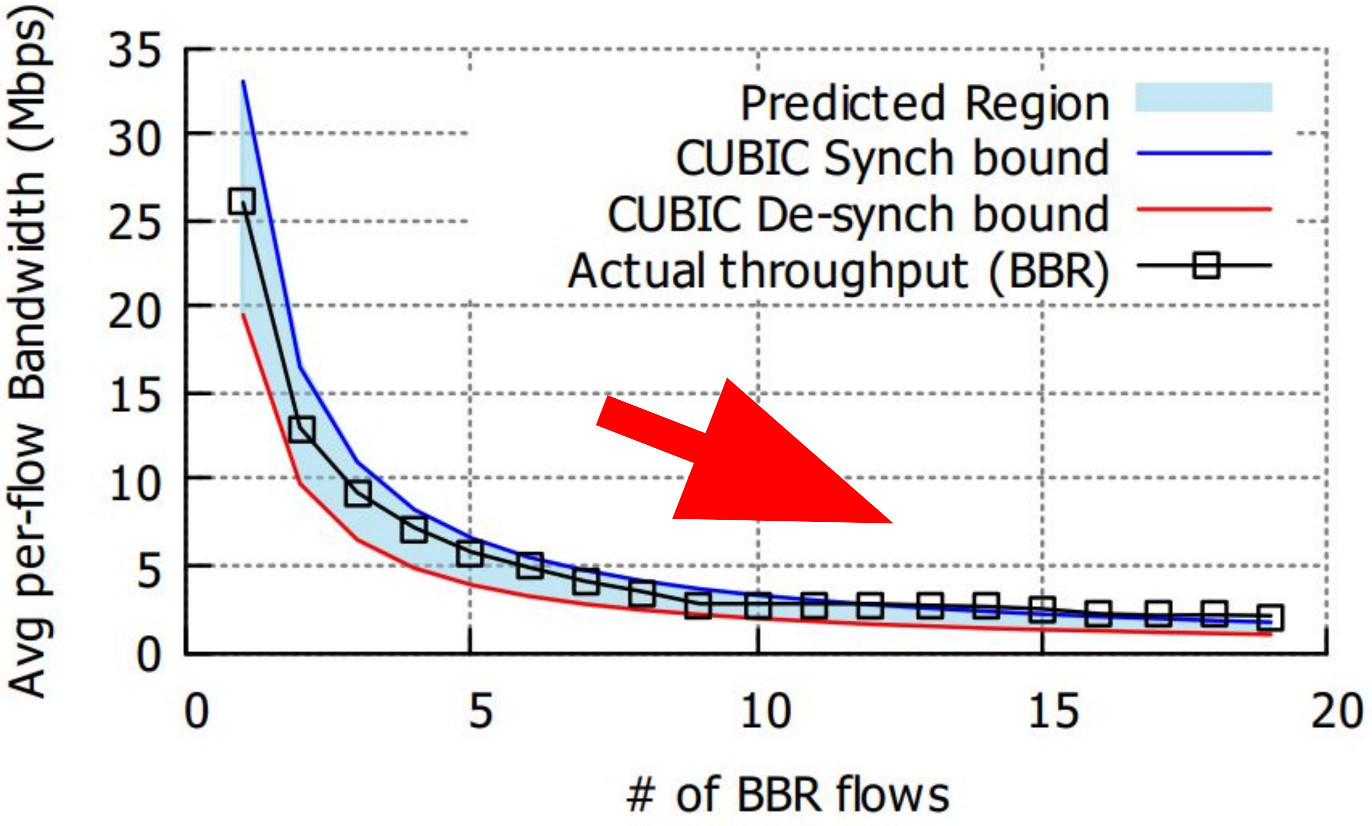
Mathematical model

BBR's throughput as more flows run BBR

Ran 20 flows through a 100 Mbps 40 ms link

Progressively increased the number of BBR flows. All other flows ran CUBIC

3 BDP buffer



Key trend:

As the number of BBR flows at the bottleneck increases, their per-flow average bandwidth decreases!

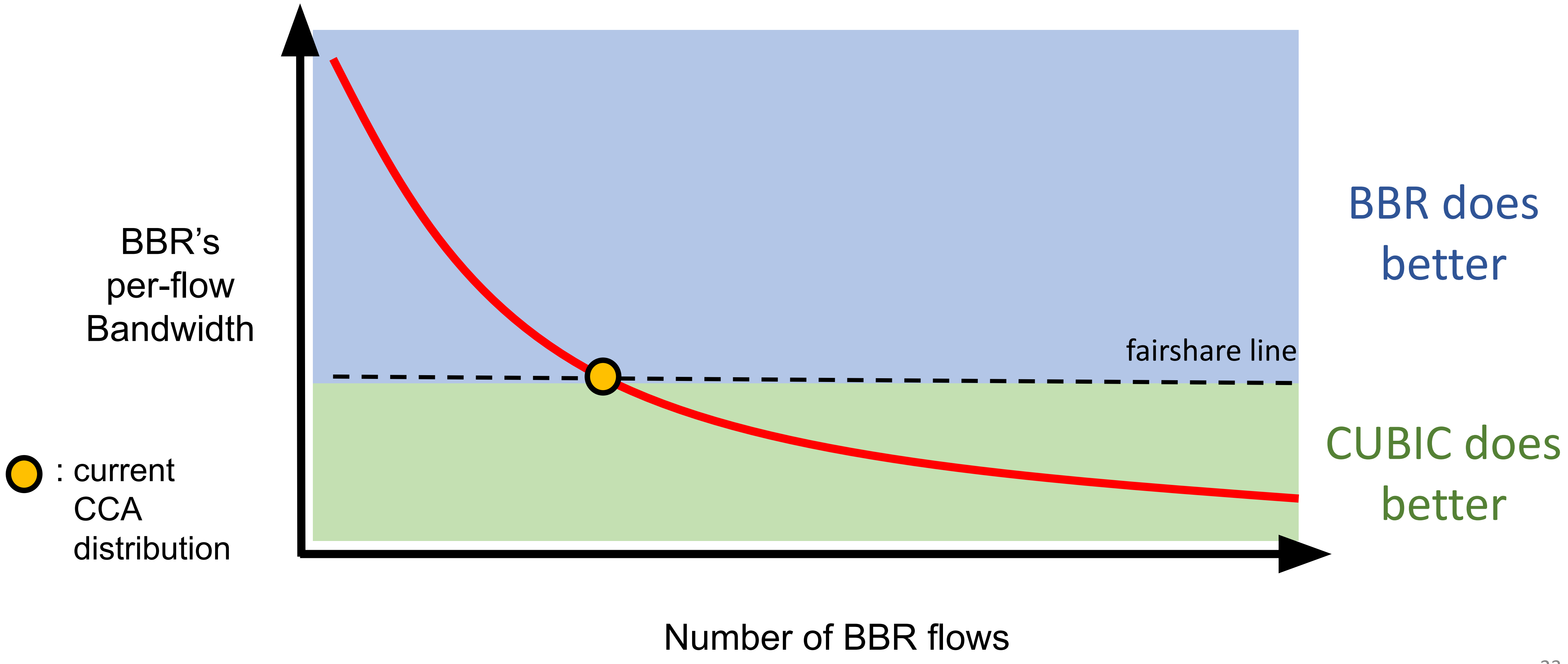
How low is too low?

Nash Equilibrium distribution of CUBIC and BBR

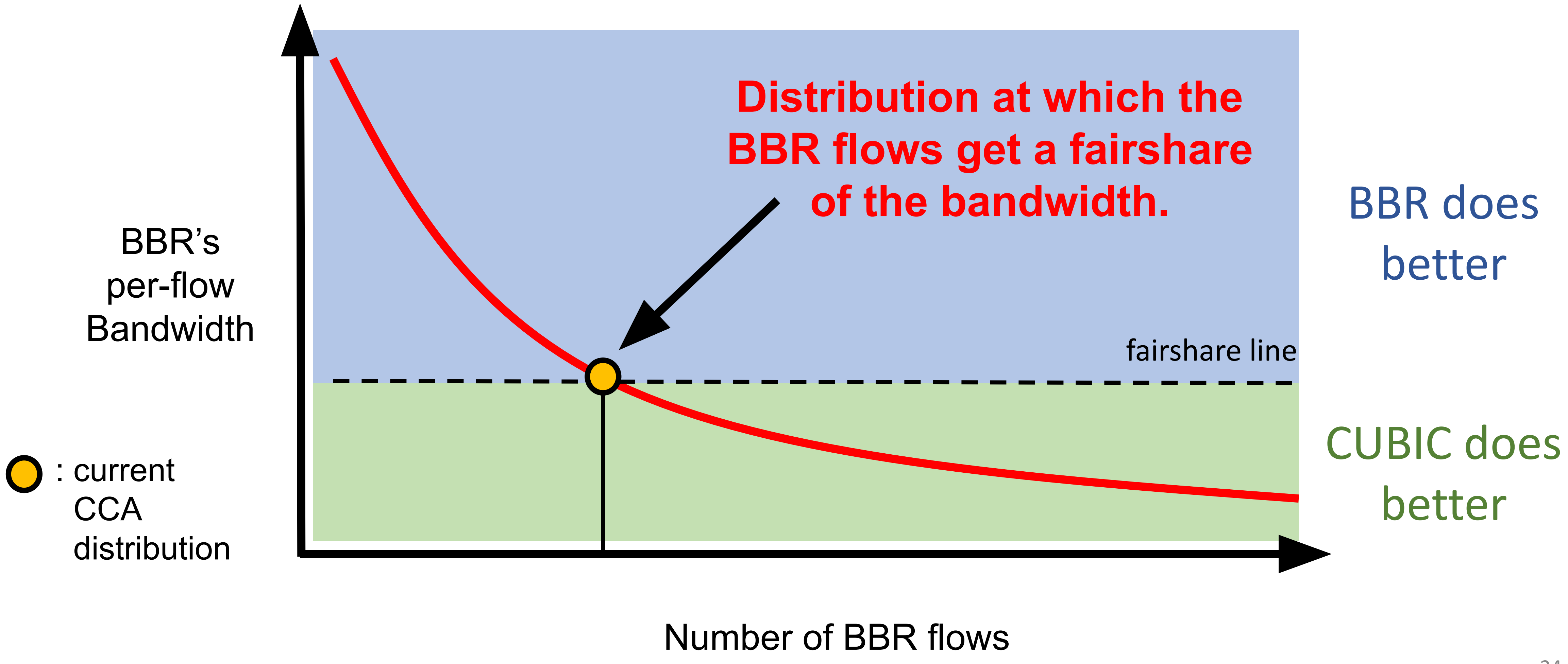
A given distribution of CUBIC and BBR flows in a network is the Nash Equilibrium (NE) if none of the flows can increase their throughput by changing algorithms.

If websites choose between CUBIC and BBR based on throughput, this is the distribution the Internet will move towards.

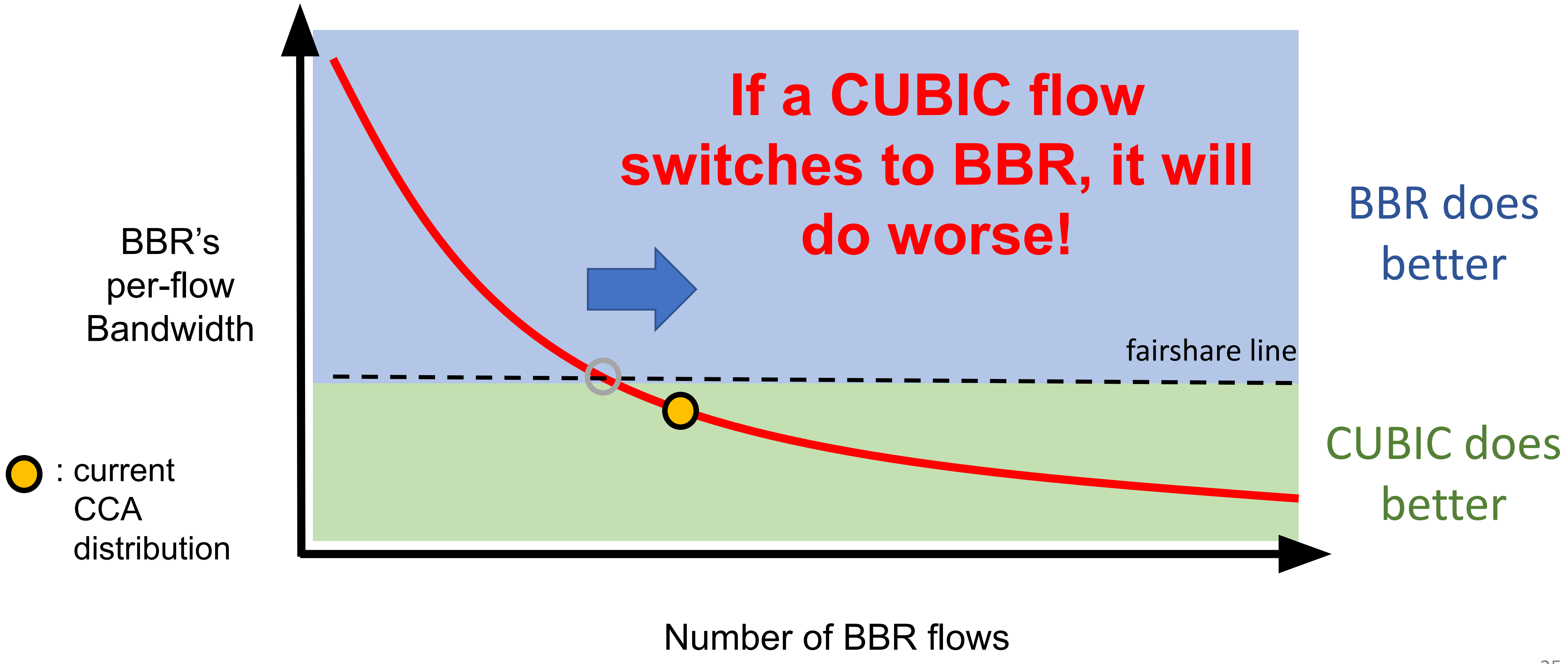
BBR's diminishing returns and the NE



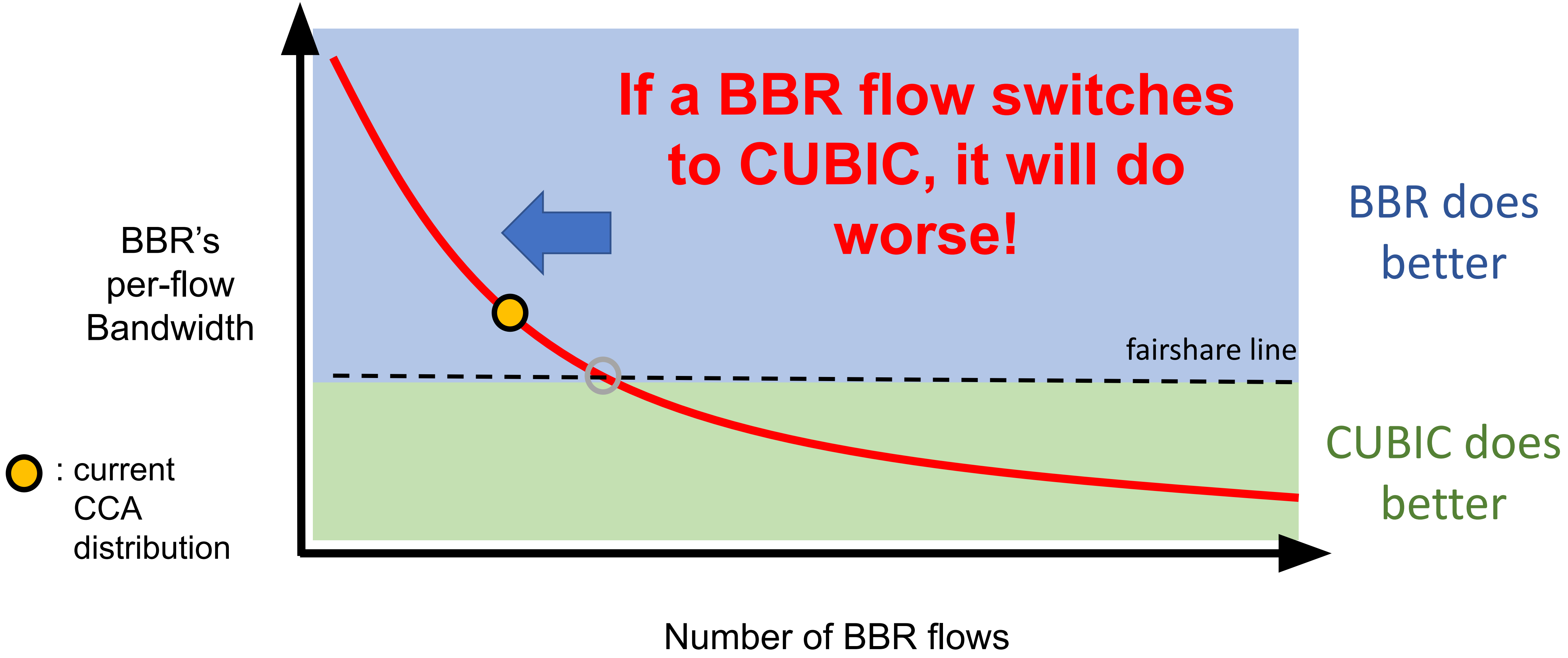
BBR's diminishing returns and the NE



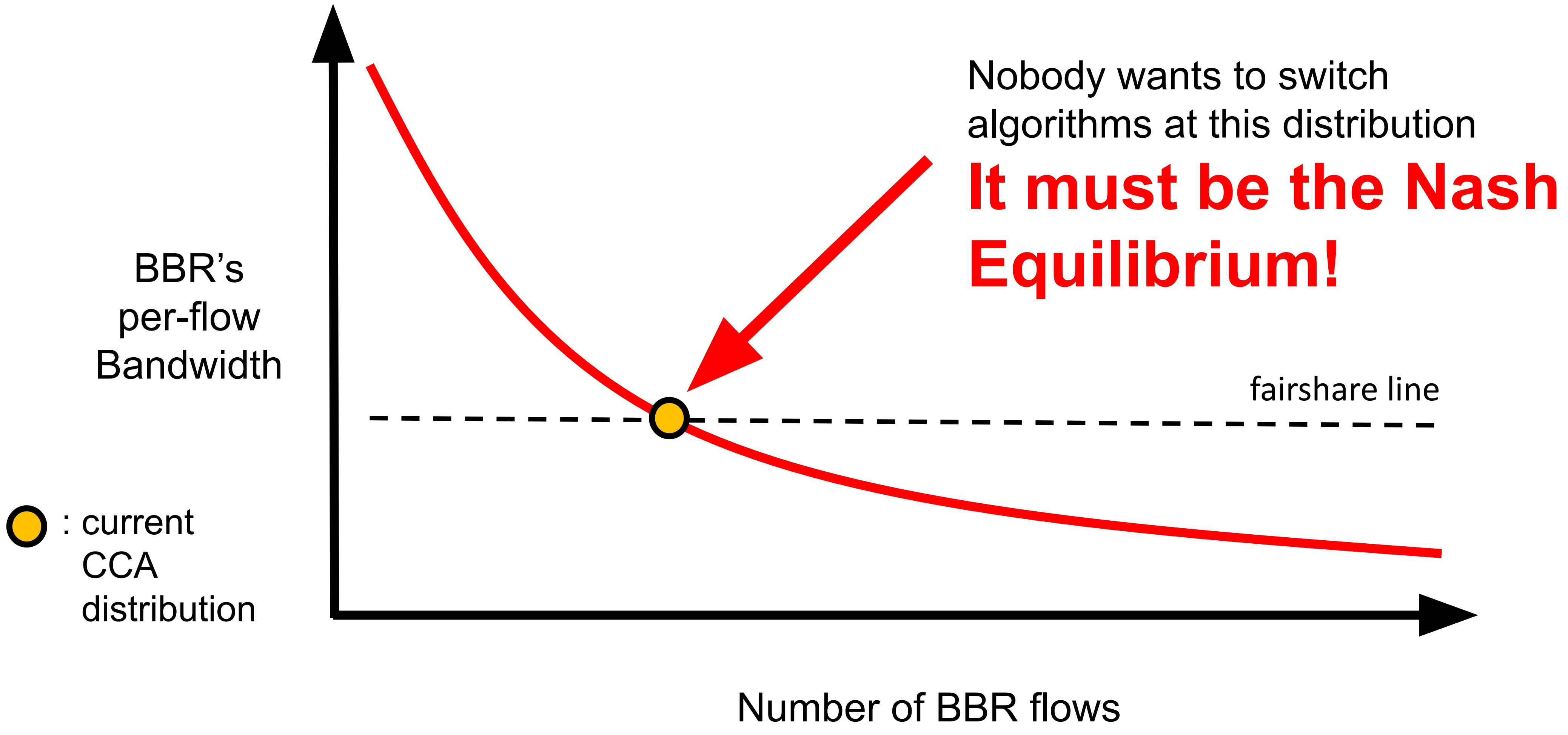
BBR's diminishing returns and the NE



BBR's diminishing returns and the NE



BBR's diminishing returns and the NE



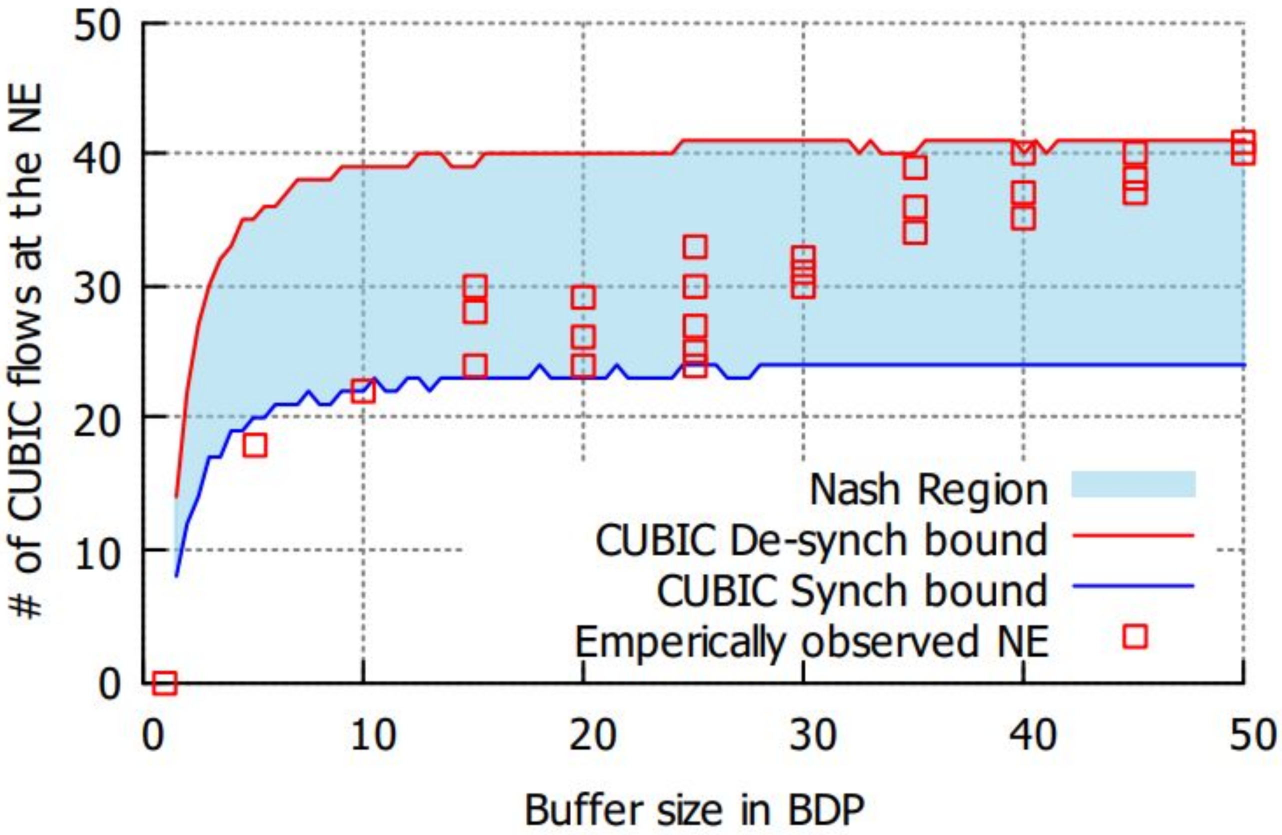
Verifying predicted Nash Equilibria

Ran 50 flows through a 50 Mbps 40 ms link

Tested all combinations of BBR and CUBIC to empirically calculate the NE distribution

Compared to model's predictions

Empirically observed NE distributions exist within our model's bounds.



Majority of NE distributions have CUBIC flows.

CUBIC is here to stay on the Internet!

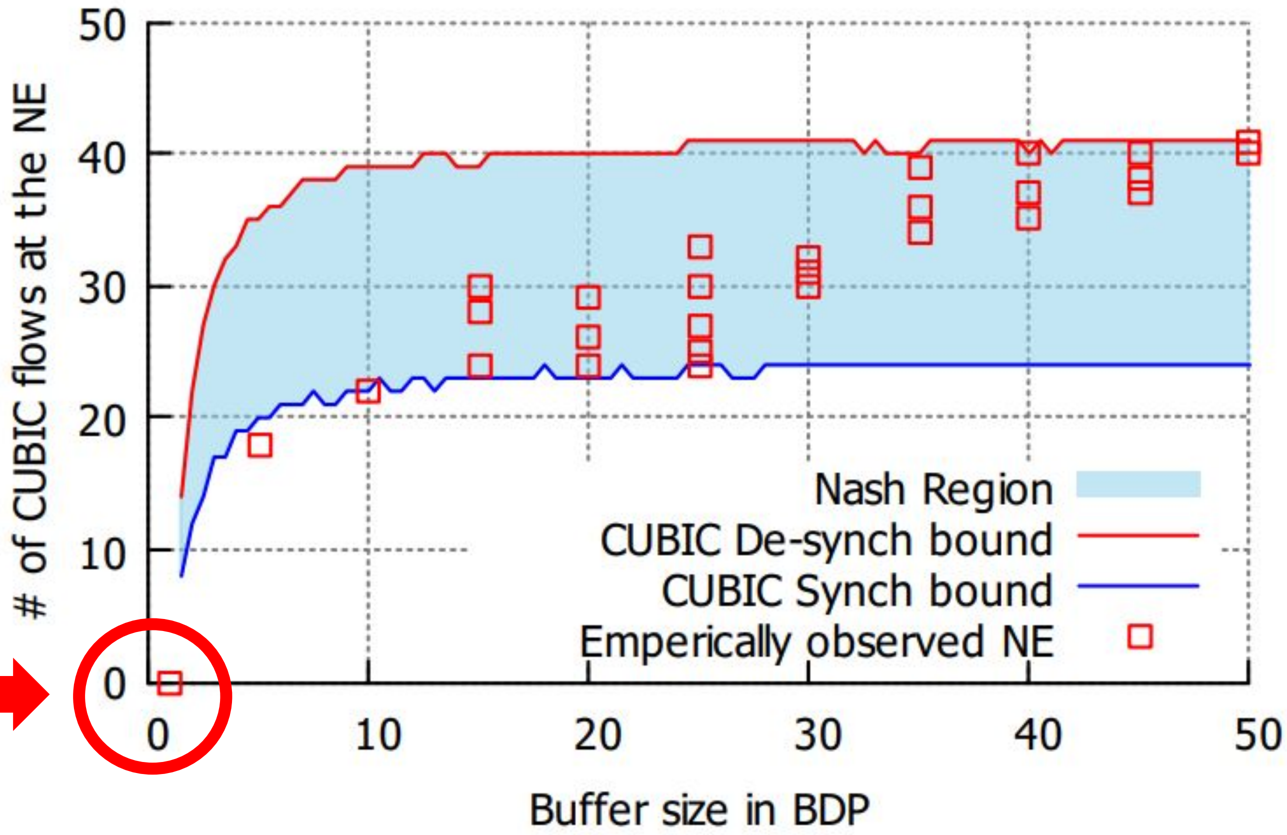
Verifying predicted Nash Equilibria

Ran 50 flows through a 50 Mbps 40 ms link

Tested all combinations of BBR and CUBIC to empirically calculate the NE distribution

Compared to model's predictions

Empirically observed NE distributions exist within our model's bounds.



Majority of NE distributions have CUBIC flows.

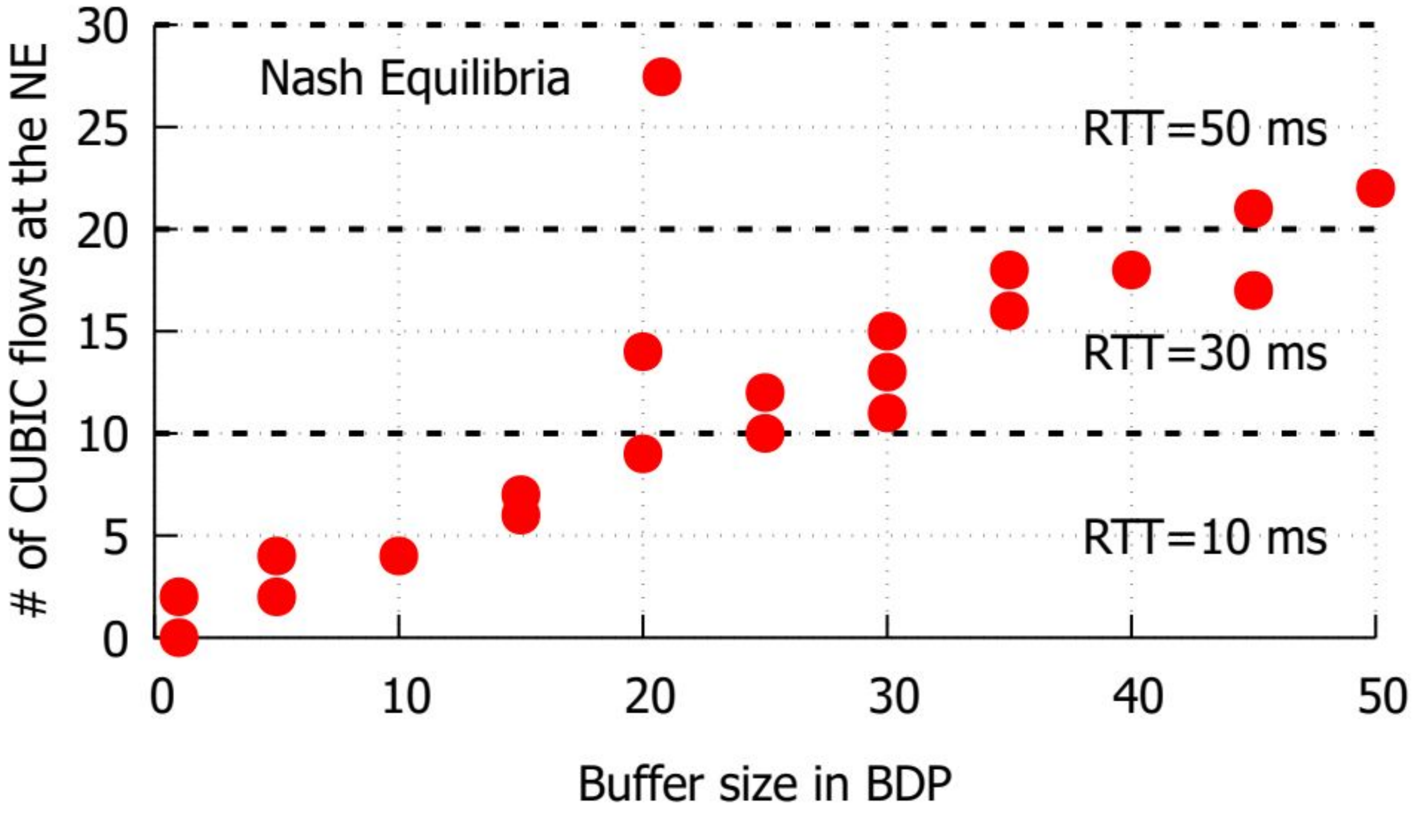
Unless the buffers are small (< 1 BDP)

CUBIC is here to stay on the Internet!

NE in Multi-RTT scenarios

Tested the model's assumption that all flows have the same RTT

NE exists for multi-RTT settings too

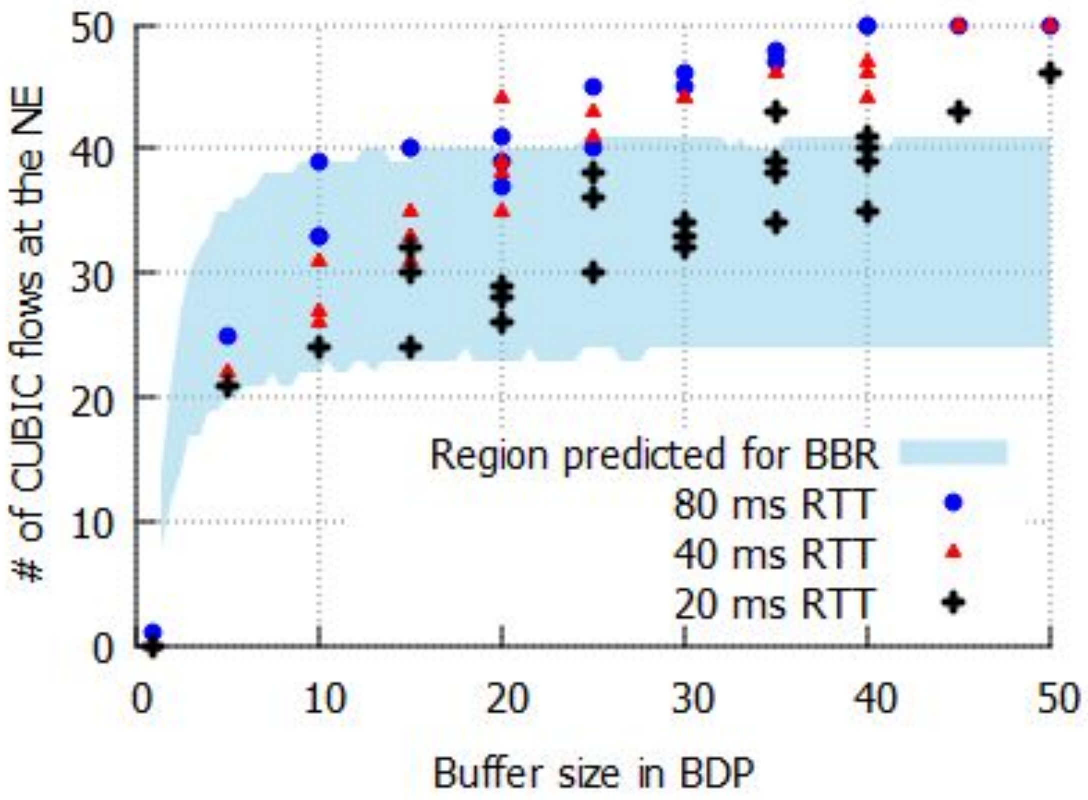


Shorter RTT flows opted for CUBIC, larger RTT flows opted for BBR at the NE

Nash Equilibria for BBRv2

Repeated experiments with BBRv2 instead of BBR

Empirically verified that mixed NE exist for BBRv2 as well



More CUBIC flows at NE when competing with BBRv2 when compared to BBRv1

Summary

We present a mathematical model for predicting the throughput shares of competing CUBIC and BBR flows.

As the number of BBR flows increases at the bottleneck, their throughput advantage will reduce.

Our game theoretic analysis shows that in most networks the Nash Equilibrium distribution of CUBIC and BBR flows will be mixed.

Context matters

(briefly discussed on the bbr dev mailing list)

This paper only explores the steady state behavior of bulk CUBIC and BBR.

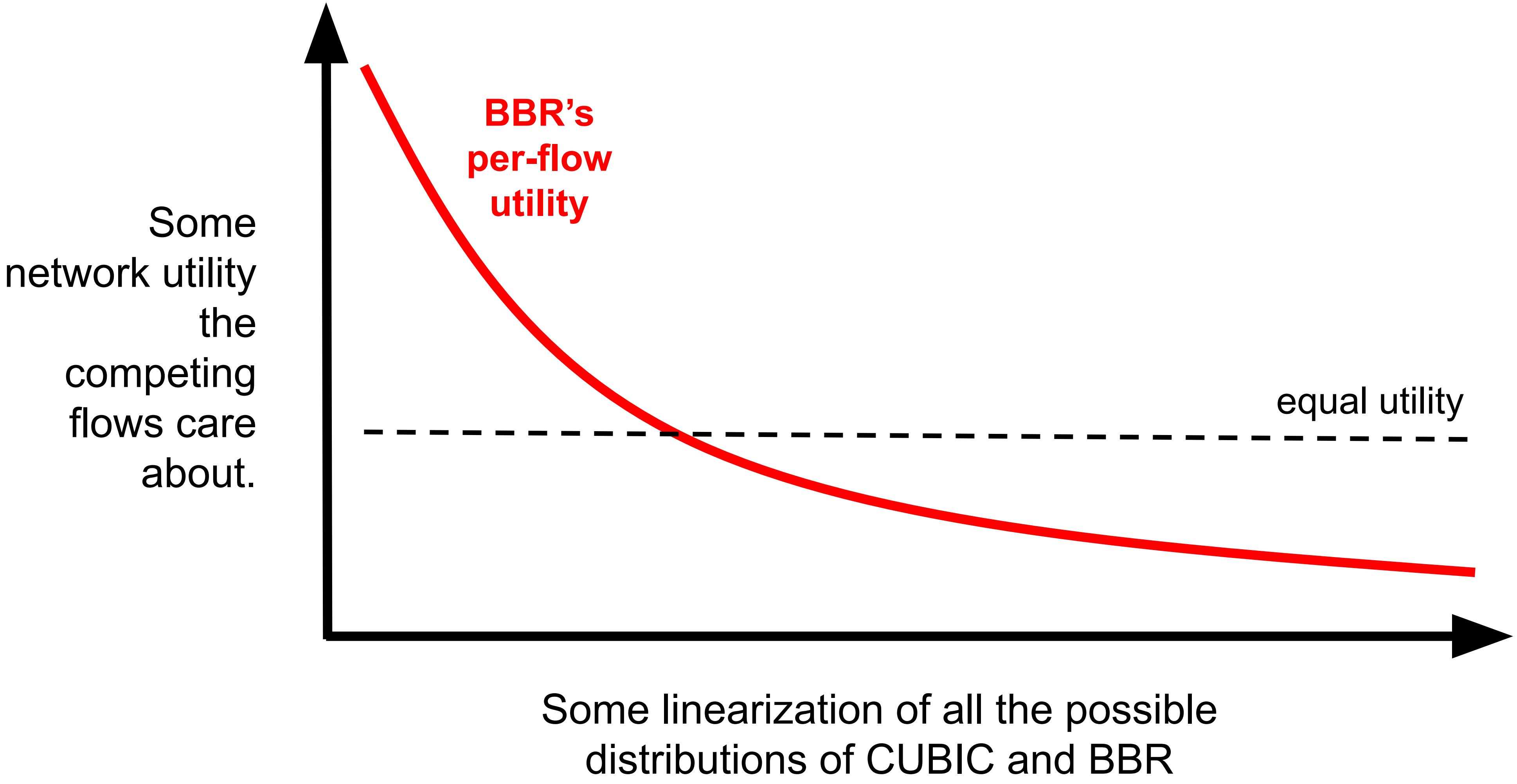
For more complex workloads with different flow sizes, we can utilize more accurate fluid models [1] that model the transient states too.

Exact NE distribution is going to depend on a variety of factors.

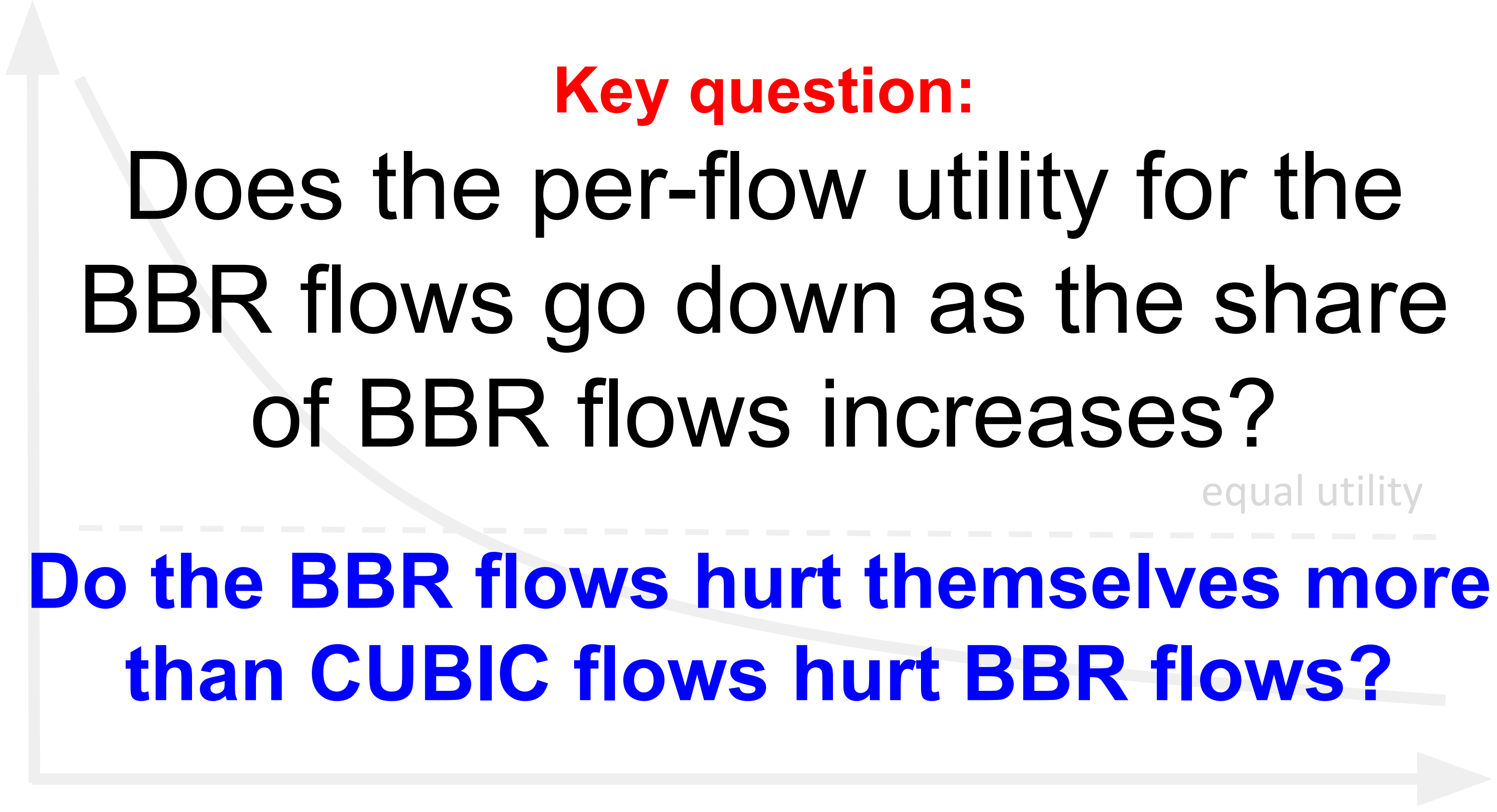
These factors include the network characteristics as well as the choice of network utility.

[1] Model-Based Insights on the Performance, Fairness, and Stability of BBR, Scherrer et al. IMC 2022

Is there a mixed NE in your network?



Is there a mixed NE in your network?



Key question:

Does the per-flow utility for the BBR flows go down as the share of BBR flows increases?

Do the BBR flows hurt themselves more than CUBIC flows hurt BBR flows?

Future Research Questions

Will a purely performance driven switch to a new congestion control algorithm ever be possible?

Taming the Zoo: How do we design for a heterogeneous congestion control landscape?

Heterogeneity in QUIC Congestion Control

Thank you!

Read the paper:

Get in touch:

ayush@comp.nus.edu.sg

Are we heading towards a BBR-dominant Internet?

Ayush Mishra, Wee Han Tiu, and Ben Leong
National University of Singapore

ABSTRACT

Since its introduction in 2016, BBR has grown in popularity rapidly and likely already accounts for more than 40% of the Internet's downstream traffic. In this paper, we investigate the following question: given BBR's performance benefits and rapid adoption, *is BBR likely to completely replace CUBIC* just like how CUBIC replaced New Reno?

We present a mathematical model that allows us to estimate BBR's throughput to within a 5% error when competing with CUBIC flows. Using this model, we show that even though BBR currently has a throughput advantage over CUBIC, this advantage will be diminished as the proportion of BBR flows increases.

Therefore, if throughput is a key consideration, it is likely that the Internet will reach a stable mixed distribution of CUBIC and BBR flows. This mixed distribution will be a *Nash Equilibrium* where none of the flows will have the performance incentive to switch between CUBIC and BBR. Our methodology is also applicable to

This is an important question because the stability of the Internet depends on the competing flows interacting well with one another. We have not experienced a *congestion collapse* [17] for many years likely because the vast majority of flows have been well-understood AIMD/MIMD-window-based TCP flows [9]. The last major change in the Internet congestion landscape happened when CUBIC replaced New Reno [22, 31]. That transition was however relatively incremental because both CUBIC and New Reno are loss-based and cwnd-based. Therefore, all existing in-network solutions, policing algorithms, and AQMs already deployed on the Internet could largely remain unchanged.

On the other hand, if BBR were to replace CUBIC as the dominant congestion control algorithm for the Internet, it represents a fundamental paradigm shift. Many classic networking questions that have supposedly been settled would have to be re-evaluated. For example, it was said that router buffers ought to be sized inversely proportional to $\sqrt{N}$, where N is the number of flows [2]. Later,