APNET 2021

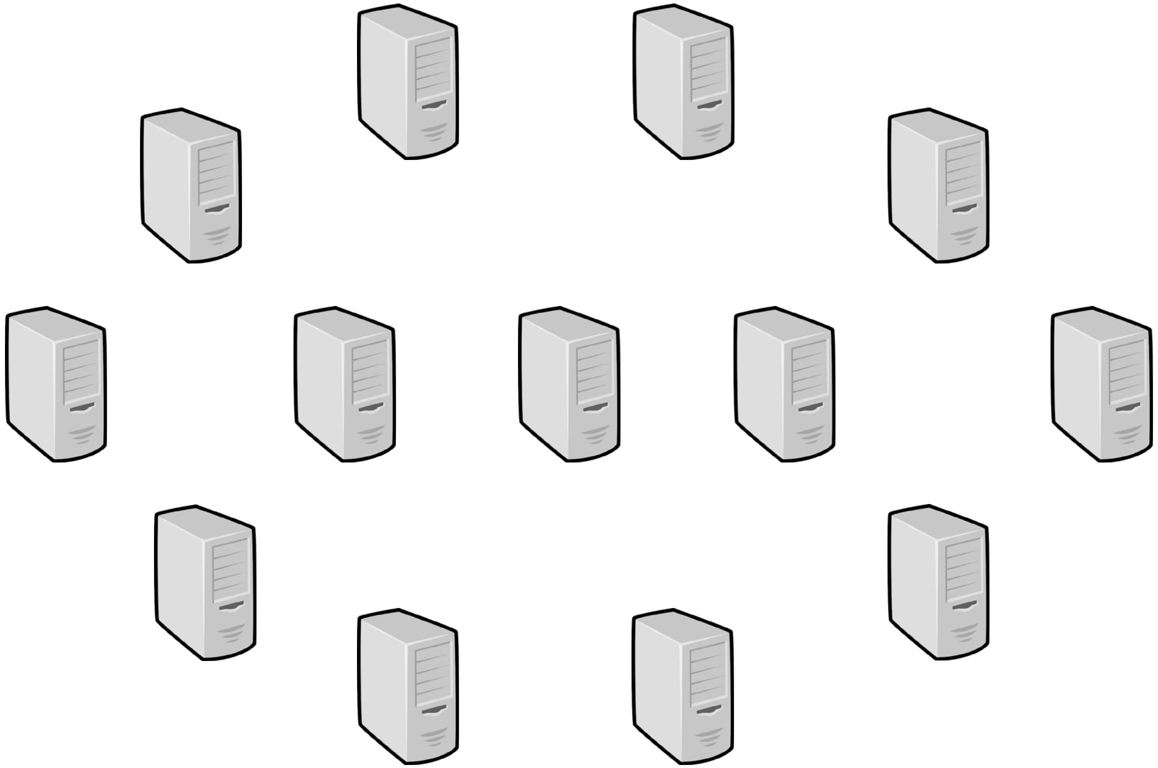
CONJECTURE:

EXISTENCE OF NASH EQUILIBRIA IN MODERN INTERNET CONGESTION CONTROL

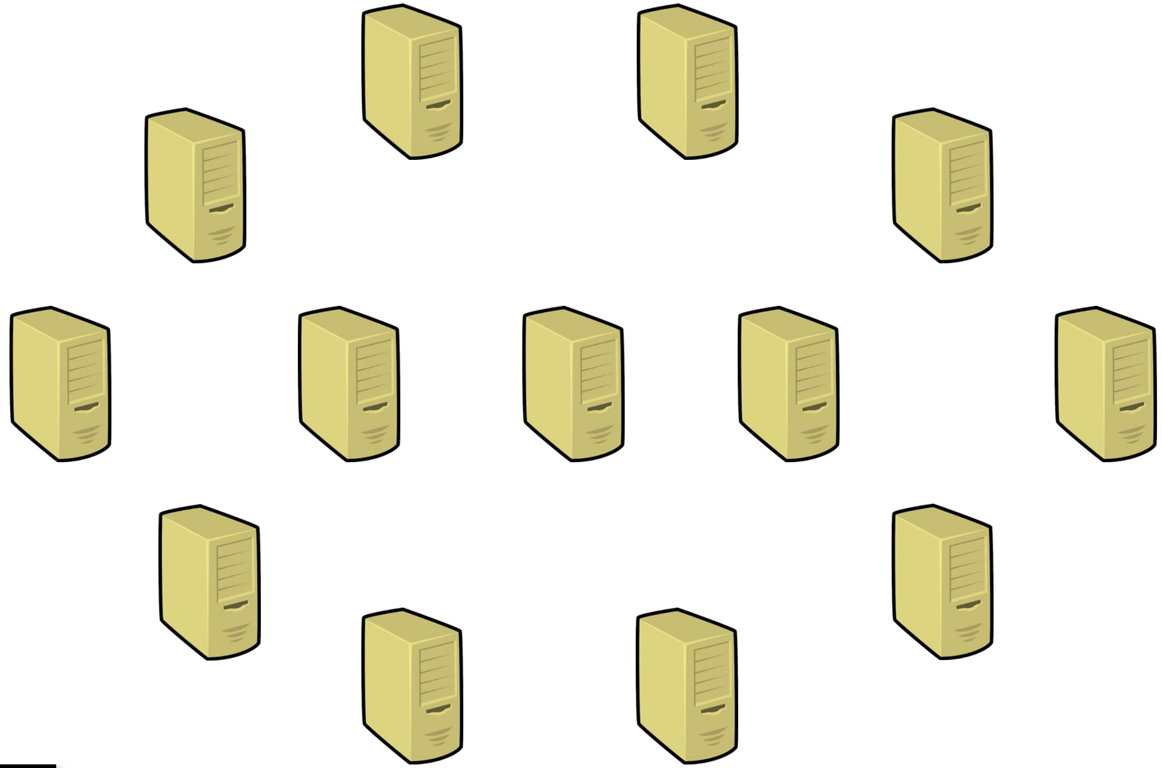
Ayush Mishra[1], Jingzhi Zhang[2], Melodies Sim[1], Sean Ng[1], Raj Joshi[1], Ben Leong[1]

[1] National University of Singapore
[2] Wuhan University

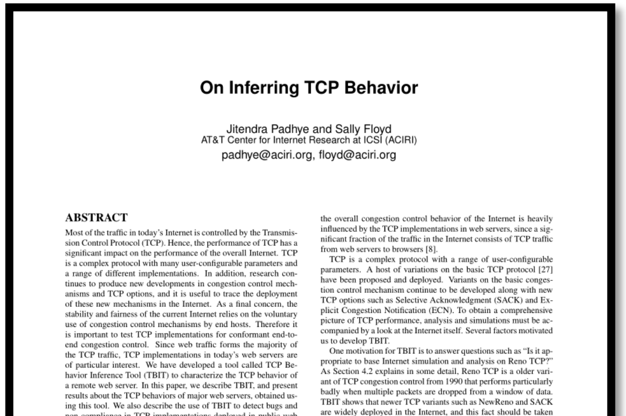
THE INTERNET



THE INTERNET



TCP Reno



On Inferring TCP Behavior

Jitendra Padhye and Sally Floyd
AT&T Center for Internet Research at ICSI (ACIRI)
padhye@aciri.org, floyd@aciri.org

ABSTRACT

Most of the traffic in today's Internet is controlled by the Transmission Control Protocol (TCP). Hence, the performance of TCP has a significant impact on the performance of the overall Internet. TCP is a complex protocol with many user-configurable parameters and a range of different implementations. In addition, research continues to produce new developments in congestion control mechanisms and TCP options, and it is useful to trace the deployment of these new mechanisms in the Internet. As a final concern, the stability and fairness of the current Internet relies on the voluntary use of congestion control mechanisms by end hosts. Therefore it is important to test TCP implementations for conformant end-to-end congestion control. Since web traffic forms the majority of the TCP traffic, TCP implementations in today's web servers are of particular interest. We have developed a tool called TCP Behavior Inference Tool (TBIT) to characterize the TCP behavior of a remote web server. In this paper, we describe TBIT, and present results about the TCP behaviors of major web servers, obtained using this tool. We also describe the use of TBIT to detect bugs and

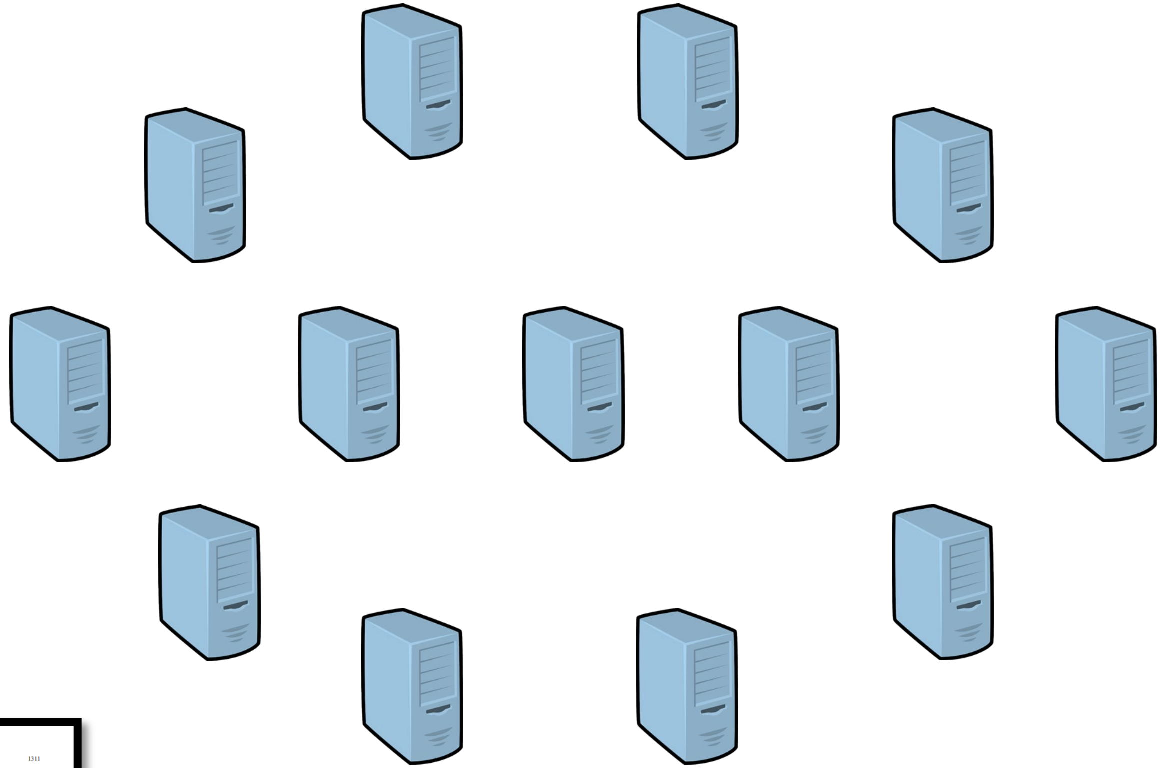
the overall congestion control behavior of the Internet is heavily influenced by the TCP implementations in web servers, since a significant fraction of the traffic in the Internet consists of TCP traffic from web servers to browsers [8].

TCP is a complex protocol with a range of user-configurable parameters. A host of variations on the basic TCP protocol [27] have been proposed and deployed. Variants on the basic congestion control mechanism continue to be developed along with new TCP options such as Selective Acknowledgment (SACK) and Explicit Congestion Notification (ECN). To obtain a comprehensive picture of TCP performance, analysis and simulations must be accompanied by a look at the Internet itself. Several factors motivated us to develop TBIT.

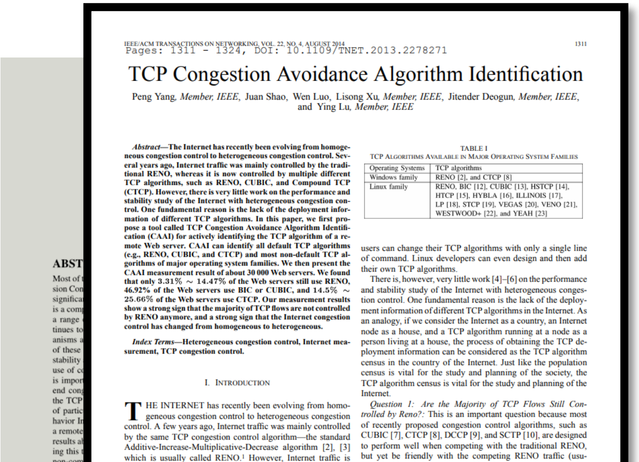
One motivation for TBIT is to answer questions such as "Is it appropriate to base Internet simulation and analysis on Reno TCP?" As Section 4.2 explains in some detail, Reno TCP is a older variant of TCP congestion control from 1990 that performs particularly badly when multiple packets are dropped from a window of data. TBIT shows that newer TCP variants such as NewReno and SACK are widely deployed in the Internet, and this fact should be taken

[2001] Padhya et al.

THE INTERNET



TCP CUBIC



IEEE/ACM TRANSACTIONS ON NETWORKING, VOL. 22, NO. 4, AUGUST 2014 1311
Pages: 1311 - 1324, DOI: 10.1109/TNET.2013.2278271

TCP Congestion Avoidance Algorithm Identification

Peng Yang, *Member, IEEE,* Juan Shao, Wen Luo, Lisong Xu, *Member, IEEE,* Jitender Deogun, *Member, IEEE,* and Ying Lu, *Member, IEEE*

***Abstract*—The Internet has recently been evolving from homogeneous congestion control to heterogeneous congestion control. Several years ago, Internet traffic was mainly controlled by the traditional RENO, whereas it is now controlled by multiple different TCP algorithms, such as RENO, CUBIC, and Compound TCP (CTCP). However, there is very little work on the performance and stability study of the Internet with heterogeneous congestion control. One fundamental reason is the lack of the deployment information of different TCP algorithms. In this paper, we first propose a tool called TCP Congestion Avoidance Algorithm Identification (CAAI) for actively identifying the TCP algorithm of a remote Web server. CAAI can identify all default TCP algorithms (e.g., RENO, CUBIC, and CTCP) and most non-default TCP algorithms of major operating system families. We then present the CAAI measurement result of about 30 000 Web servers. We found that only 3.31% ~ 14.47% of the Web servers still use RENO, 46.92% of the Web servers use BIC or CUBIC, and 14.5% ~ 25.66% of the Web servers use CTCP. Our measurement results show a strong sign that the majority of TCP flows are not controlled by RENO anymore, and a strong sign that the Internet congestion control has changed from homogeneous to heterogeneous.**

***Index Terms*—Heterogeneous congestion control, Internet measurement, TCP congestion control.**

I. Introduction

THE INTERNET has recently been evolving from homogeneous congestion control to heterogeneous congestion control. A few years ago, Internet traffic was mainly controlled by the same TCP congestion control algorithm—the standard Additive-Increase-Multiplicative-Decrease algorithm [2], [3] which is usually called RENO.[1] However, Internet traffic is

TABLE I
TCP Algorithms Available in Major Operating System Families

Operating Systems	TCP algorithms
Windows family	RENO [2], and CTCP [8]
Linux family	RENO, BIC [12], CUBIC [13], HSTCP [14], HTCP [15], HYBLA [16], ILLINOIS [17], LP [18], STCP [19], VEGAS [20], VENO [21], WESTWOOD+ [22], and YEAH [23]

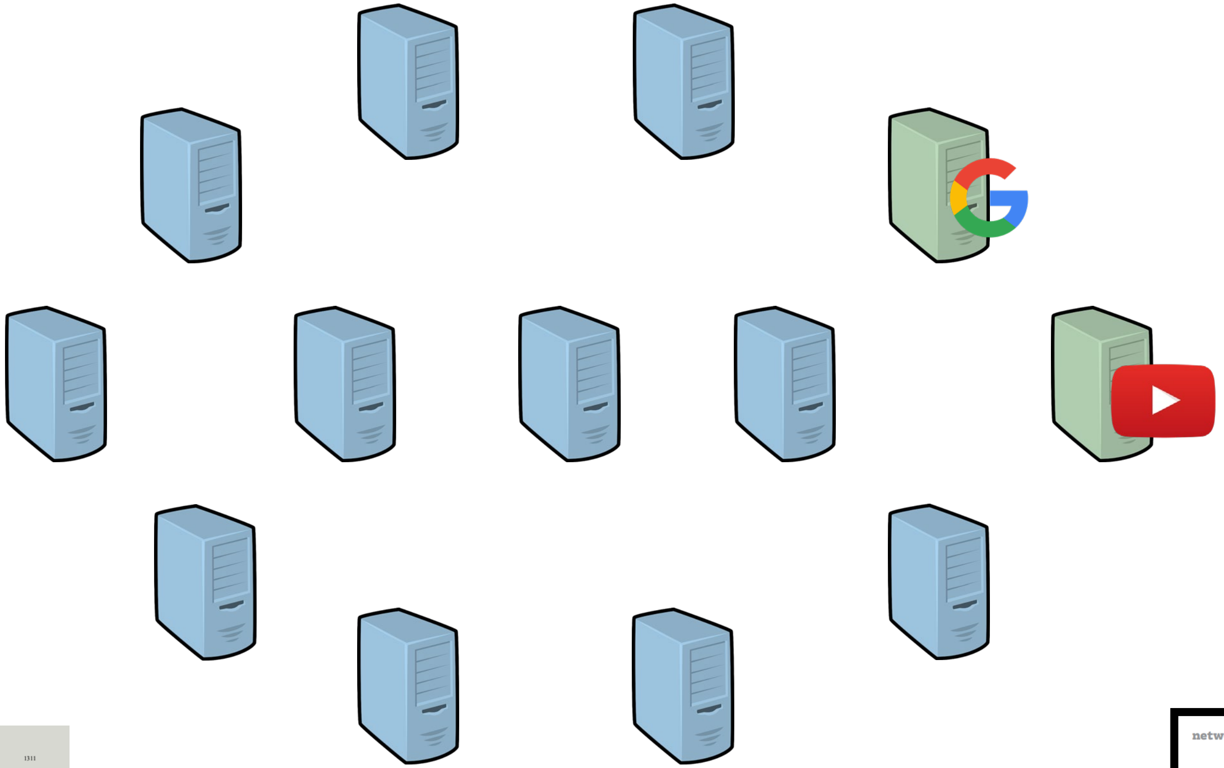
users can change their TCP algorithms with only a single line of command. Linux developers can even design and then add their own TCP algorithms.

There is, however, very little work [4]–[6] on the performance and stability study of the Internet with heterogeneous congestion control. One fundamental reason is the lack of the deployment information of different TCP algorithms in the Internet. As an analogy, if we consider the Internet as a country, an Internet node as a house, and a TCP algorithm running at a node as a person living at a house, the process of obtaining the TCP deployment information can be considered as the TCP algorithm census in the country of the Internet. Just like the population census is vital for the study and planning of the society, the TCP algorithm census is vital for the study and planning of the Internet.

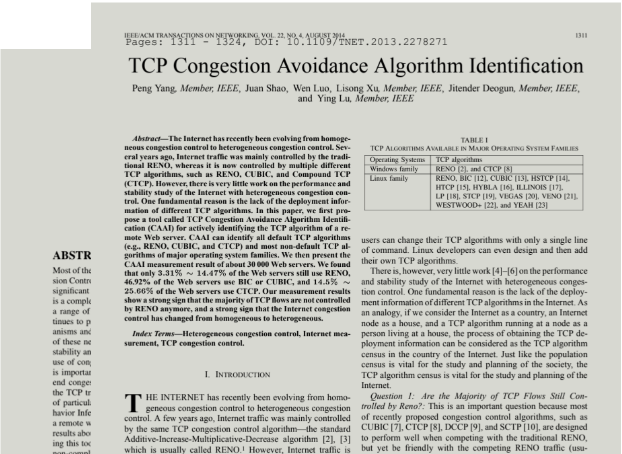
Question 1: Are the Majority of TCP Flows Still Controlled by Reno?: This is an important question because most of recently proposed congestion control algorithms, such as CUBIC [7], CTCP [8], DCCP [9], and SCTP [10], are designed to perform well when competing with the traditional RENO, but yet be friendly with the competing RENO traffic (usu-

[2011] Yang et al.

THE INTERNET



TCP CUBIC



IEEE/ACM TRANSACTIONS ON NETWORKING, VOL. 22, NO. 4, AUGUST 2014

Pages: 1311 – 1324, DOI: 10.1109/TNET.2013.2278271

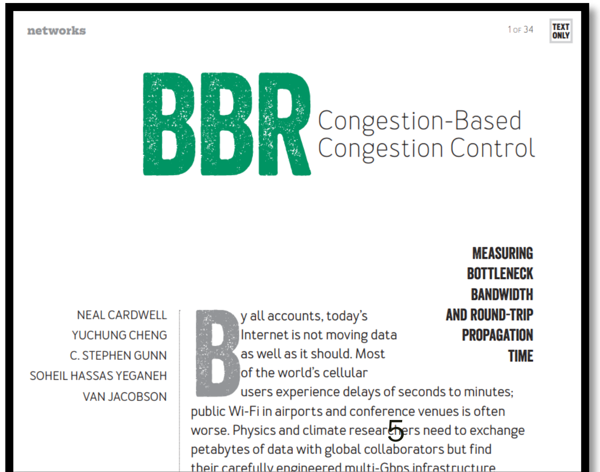
TCP Congestion Avoidance Algorithm Identification

Peng Yang, Member, IEEE, Juan Shao, Wen Luo, Lisong Xu, Member, IEEE, Jitender Deogun, Member, IEEE, and Ying Lu, Member, IEEE

[2011] Yang et al.

[2016] Cardwell et al.

TCP BBR



networks

BBR Congestion-Based Congestion Control

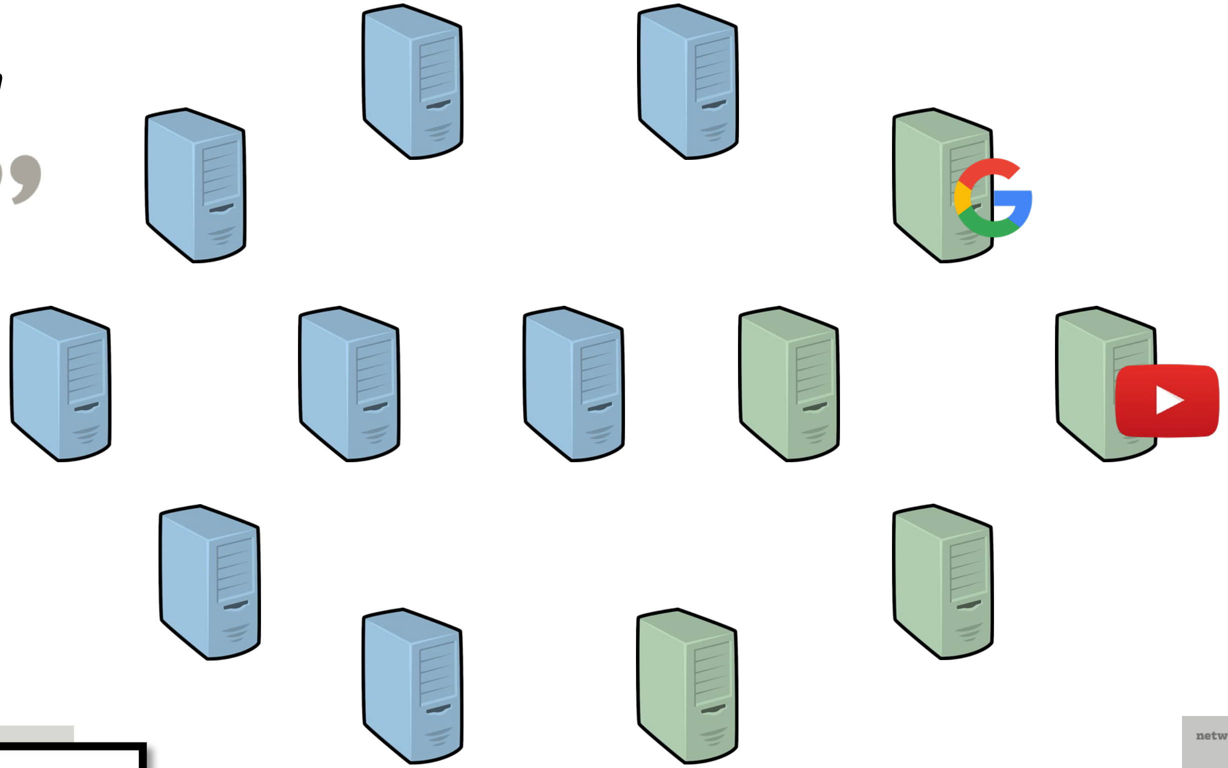
MEASURING BOTTLENECK BANDWIDTH AND ROUND-TRIP PROPAGATION TIME

NEAL CARDWELL
YUCHUNG CHENG
C. STEPHEN GUNN
SOHEIL HASSAS YEGANEH
VAN JACOBSON

By all accounts, today's Internet is not moving data as well as it should. Most of the world's cellular users experience delays of seconds to minutes; public Wi-Fi in airports and conference venues is often worse. Physics and climate researchers need to exchange petabytes of data with global collaborators but find

THE INTERNET

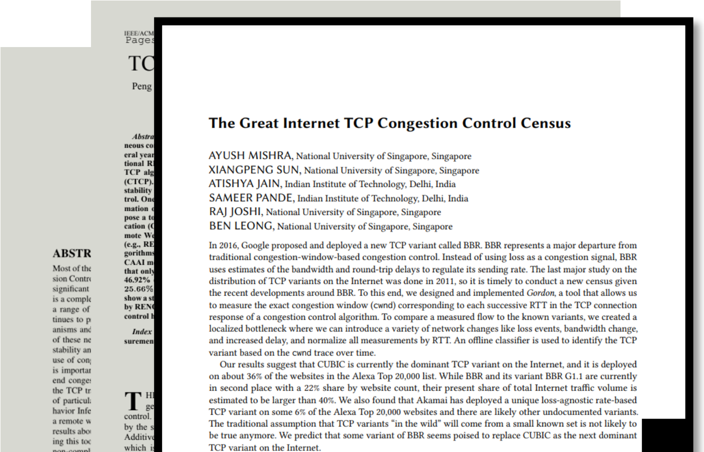
"CUBIC (36%), BBR (22%) are the most dominant congestion control algorithms on the Internet"



 TCP CUBIC

 TCP BBR

[2019] Mishra et al.

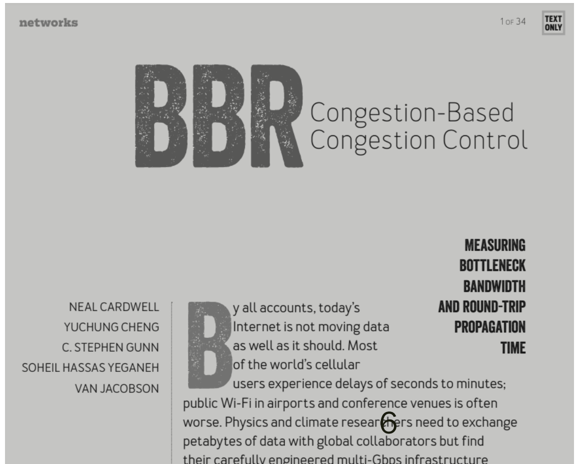


The Great Internet TCP Congestion Control Census

AYUSH MISHRA, National University of Singapore, Singapore
XIANGPENG SUN, National University of Singapore, Singapore
ATISHYA JAIN, Indian Institute of Technology, Delhi, India
SAMEER PANDE, Indian Institute of Technology, Delhi, India
RAJ JOSHI, National University of Singapore, Singapore
BEN LEONG, National University of Singapore, Singapore

In 2016, Google proposed and deployed a new TCP variant called BBR. BBR represents a major departure from traditional congestion-window-based congestion control. Instead of using loss as a congestion signal, BBR uses estimates of the bandwidth and round-trip delays to regulate its sending rate. The last major study on the distribution of TCP variants on the Internet was done in 2011, so it is timely to conduct a new census given the recent developments around BBR. To this end, we designed and implemented *Gordon*, a tool that allows us to measure the exact congestion window (cwnd) corresponding to each successive RTT in the TCP connection response of a congestion control algorithm. To compare a measured flow to the known variants, we created a localized bottleneck where we can introduce a variety of network changes like loss events, bandwidth change, and increased delay, and normalize all measurements by RTT. An offline classifier is used to identify the TCP variant based on the cwnd trace over time.

Our results suggest that CUBIC is currently the dominant TCP variant on the Internet, and it is deployed on about 36% of the websites in the Alexa Top 20,000 list. While BBR and its variant BBR G1.1 are currently in second place with a 22% share by website count, their present share of total Internet traffic volume is estimated to be larger than 40%. We also found that Akamai has deployed a unique loss-agnostic rate-based TCP variant on some 6% of the Alexa Top 20,000 websites and there are likely other undocumented variants. The traditional assumption that TCP variants "in the wild" will come from a small known set is not likely to be true anymore. We predict that some variant of BBR seems poised to replace CUBIC as the next dominant TCP variant on the Internet.



networks

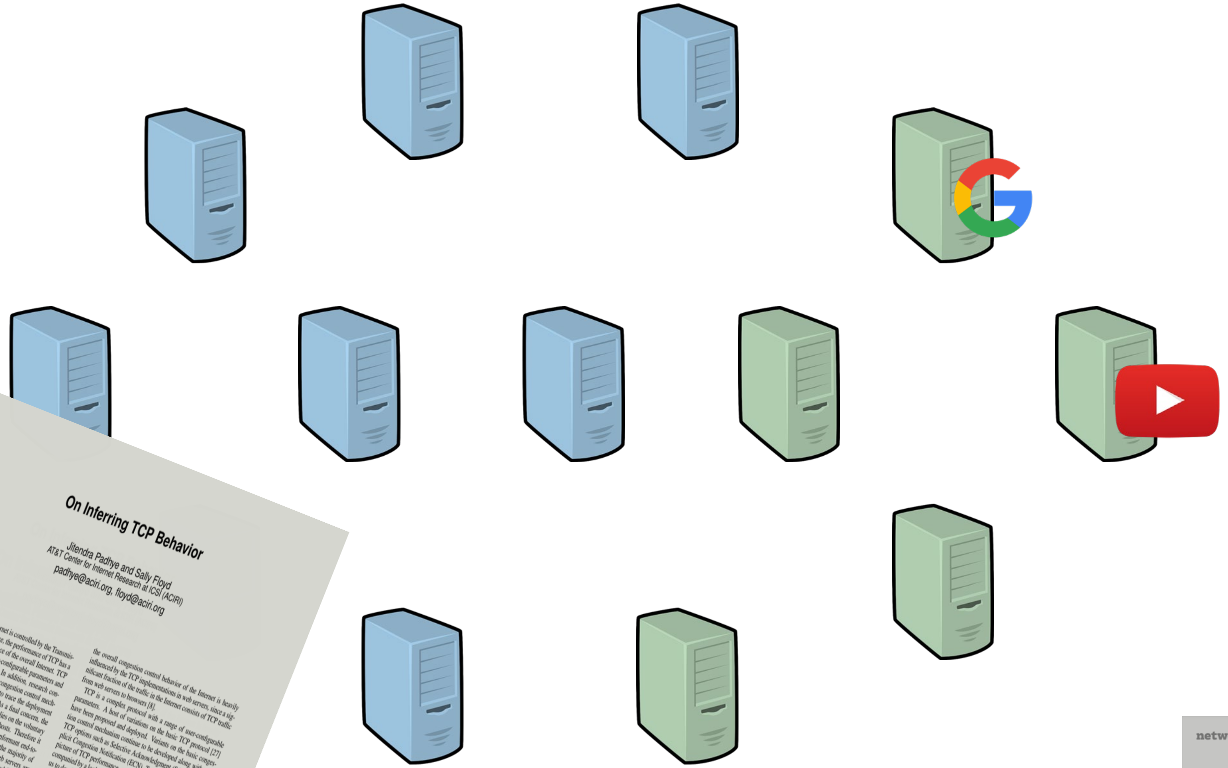
BBR Congestion-Based Congestion Control

MEASURING BOTTLENECK BANDWIDTH AND ROUND-TRIP PROPAGATION TIME

NEAL CARDWELL
YUCHUNG CHENG
C. STEPHEN GUNN
SOHEIL HASSAS YEGANEH
VAN JACOBSON

By all accounts, today's Internet is not moving data as well as it should. Most of the world's cellular users experience delays of seconds to minutes; public Wi-Fi in airports and conference venues is often worse. Physics and climate researchers need to exchange petabytes of data with global collaborators but find their carefully engineered multi-Gbps infrastructure

THE INTERNET



TC

TCP BBR


On Inferring TCP Behavior

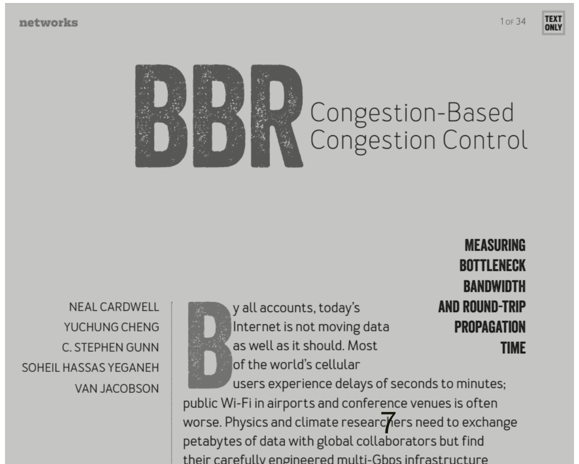
Jitendra Padhye and Sally Floyd
AT&T Center for Internet Research at ICSI (ACIRI)
padhye@aciri.org, floyd@aciri.org

ABSTRACT

be true anymore. We predict that some variant of BBR seems poised to replace CUBIC as the next dominant TCP variant on the Internet.

[2019] Mishra et al.

> *CUBIC (36%), BBR (22%) are the most dominant congestion control algorithms on the Internet*


networks

1 of 34

TEXT ONLY

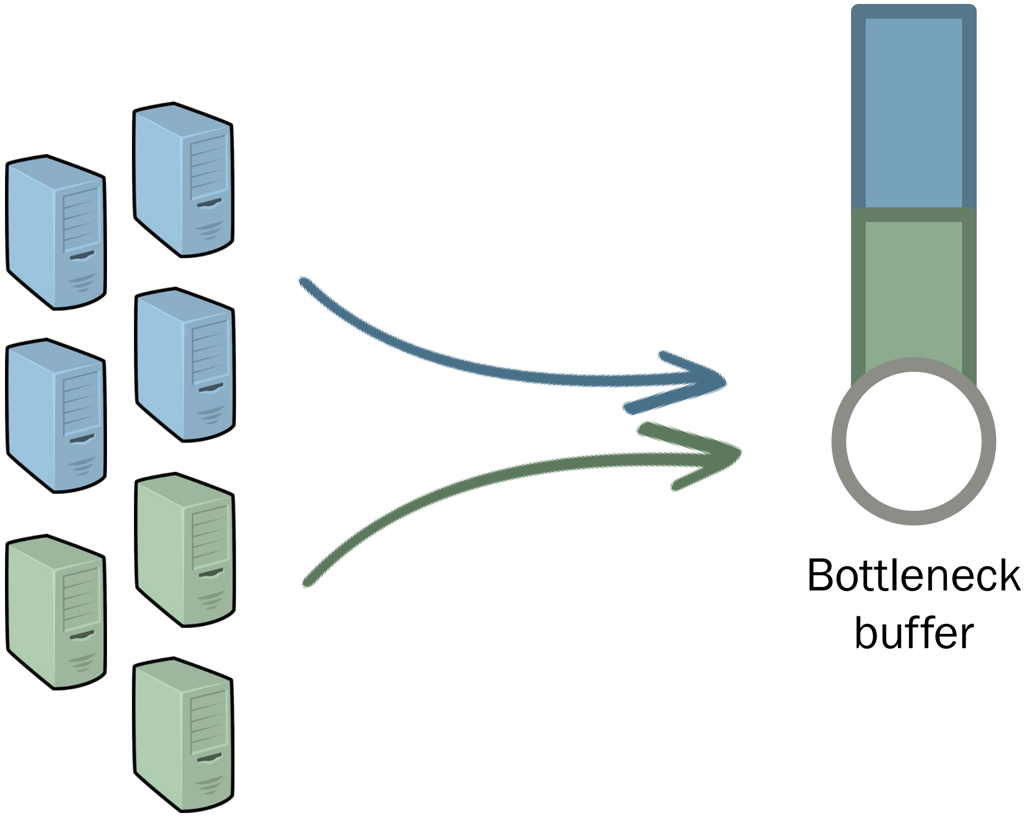
BBR Congestion-Based Congestion Control

MEASURING BOTTLENECK BANDWIDTH AND ROUND-TRIP PROPAGATION TIME

NEAL CARDWELL
YUCHUNG CHENG
C. STEPHEN GUNN
SOHEIL HASSAS YEGANEH
VAN JACOBSON

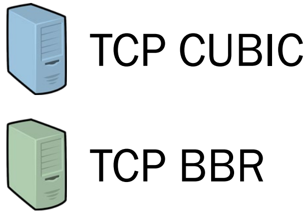
By all accounts, today's Internet is not moving data as well as it should. Most of the world's cellular users experience delays of seconds to minutes; public Wi-Fi in airports and conference venues is often worse. Physics and climate researchers need to exchange petabytes of data with global collaborators but find their carefully engineered multi-Gbps infrastructure

A SMALL NUMBER OF BBR FLOWS CAN BE VERY COMPETITIVE AGAINST CUBIC

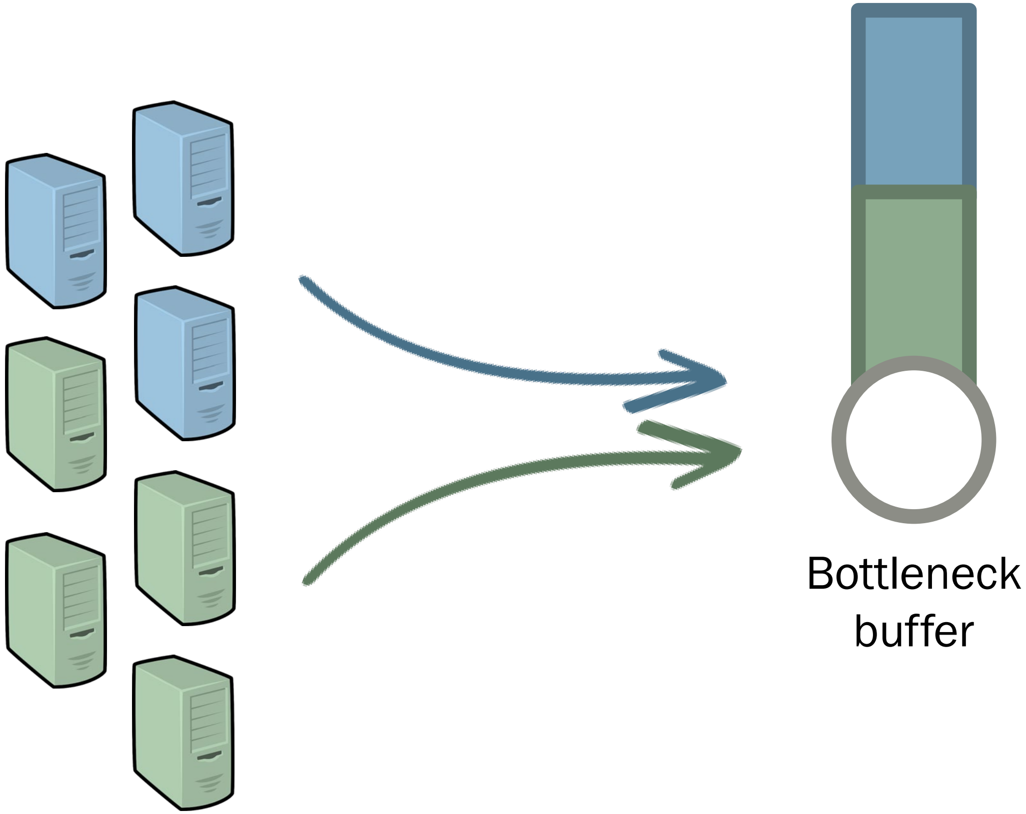


Most early deployment results from Google, Dropbox and Spotify cite better throughput as a reason to switch to BBR.

How will these benefits sustain as more and more people catch on?

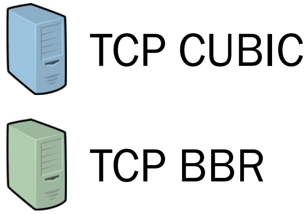


A SMALL NUMBER OF BBR FLOWS CAN BE VERY COMPETITIVE AGAINST CUBIC

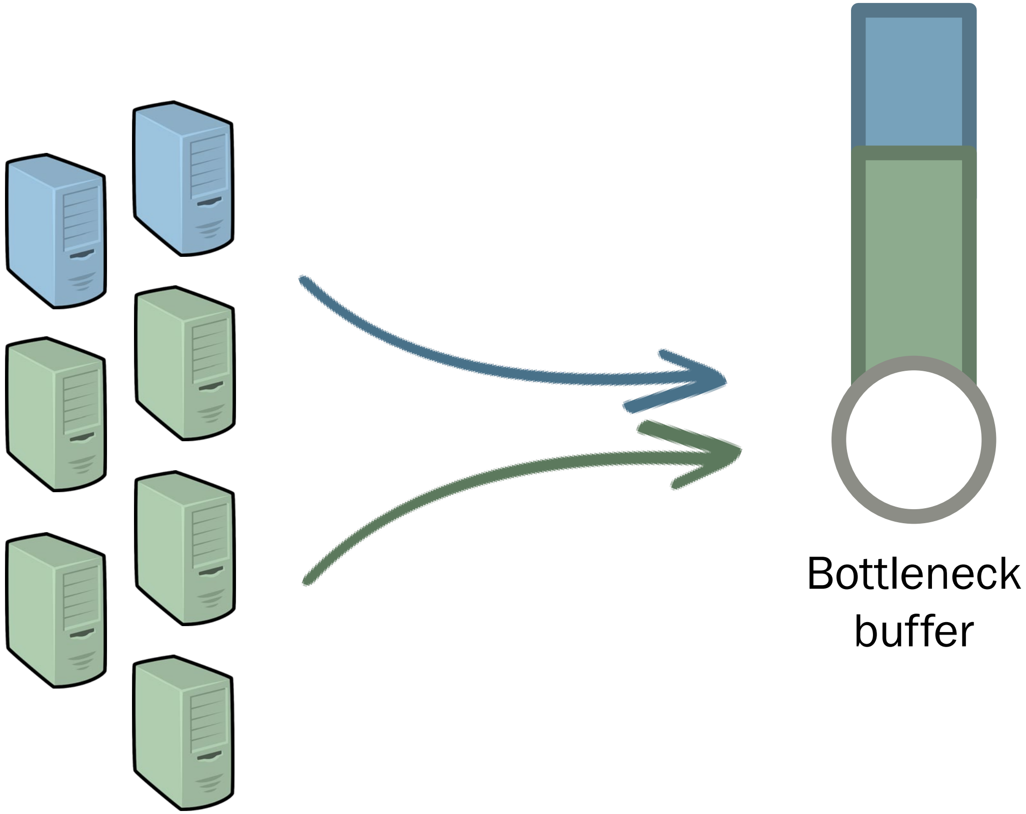


Most early deployment results from Google, Dropbox and Spotify cite better throughput as a reason to switch to BBR.

How will these benefits sustain as more and more people catch on?

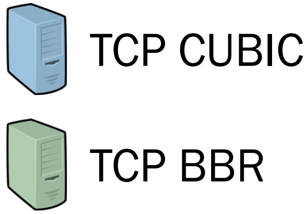


A SMALL NUMBER OF BBR FLOWS CAN BE VERY COMPETITIVE AGAINST CUBIC



Most early deployment results from Google, Dropbox and Spotify cite better throughput as a reason to switch to BBR.

How will these benefits sustain as more and more people catch on?

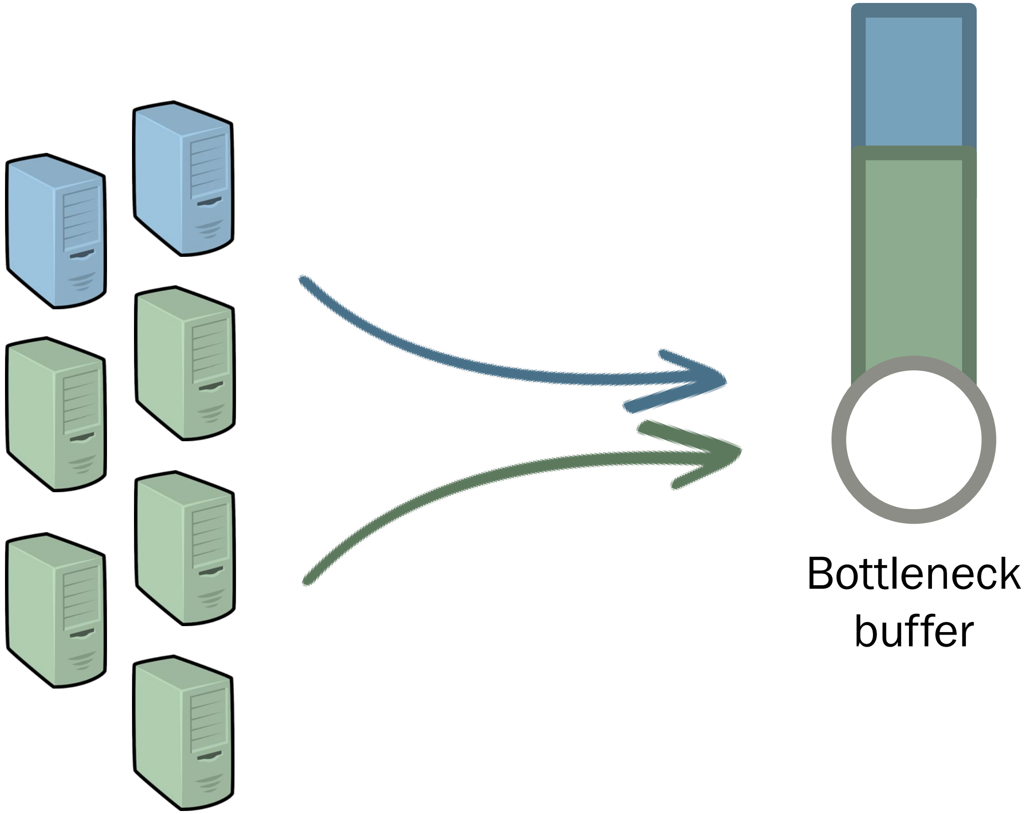


Problem Statement:

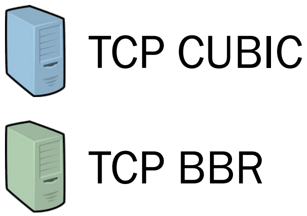
HOW DO WE EXPECT THE INTERNET TO EVOLVE?

Would it be reasonable to expect the whole Internet to switch to BBR or its variants, in the near future?

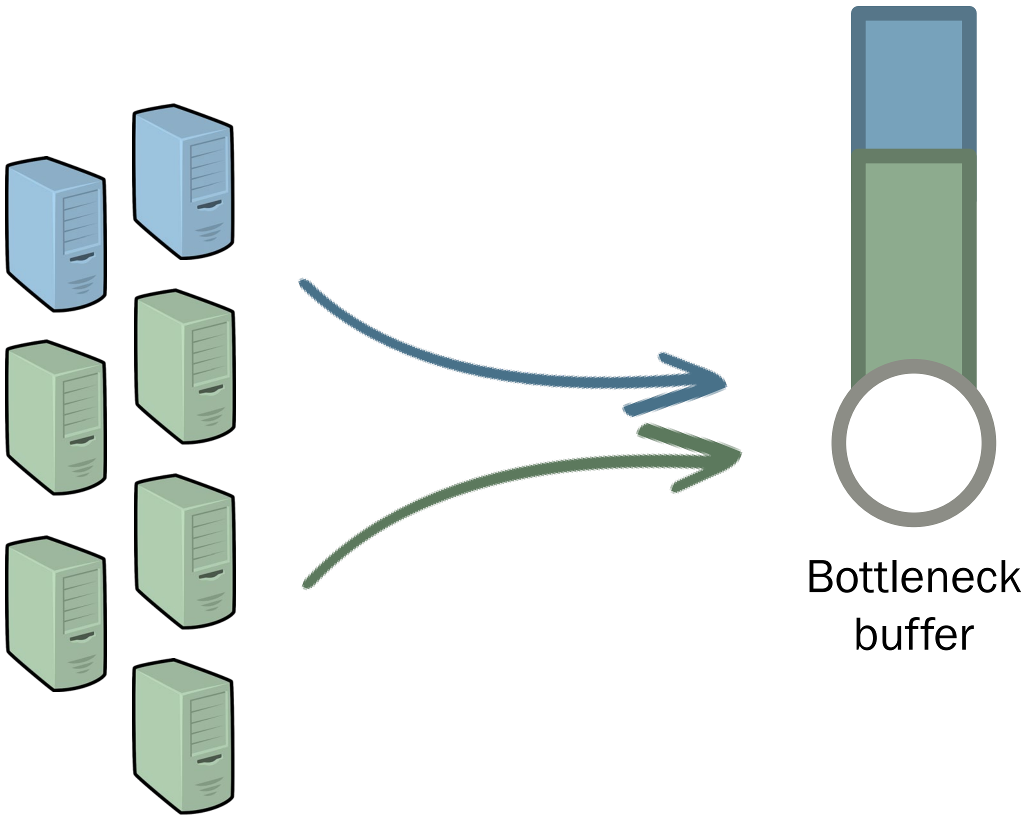
We can model the Internet as a normal form game!



In game theory, a normal form game is a standard representation of a game where the players have some preference of outcome, which in this case is throughput



We can model the Internet as a normal form game!



In game theory, a normal form game is a standard representation of a game where the players have some preference of outcome, which in this case is throughput


“Players”

Strategies:


TCP CUBIC


TCP BBR

Known Interactions between CUBIC and BBR

Throughput in relation to buffer size

- Since CUBIC is loss based, it is able to outperform BBR in deep buffers by placing a lot of packets in the bottleneck buffer.
- On the other hand, in shallow buffers, CUBIC loses out to BBR because of frequent packet losses.

Observation 1. When competing at the bottleneck where the buffer is deep, CUBIC tends to have higher throughput than BBR; the converse is true when the buffer is shallow.

Throughput in relation to buffer size

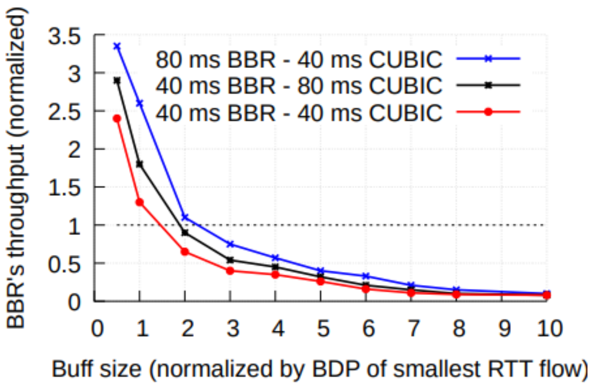


Fig. 1. BBR throughput vs. bottleneck buffer size.

- Therefore, there must be some buffer size T_{fair} where the throughput of the two competing flows must be the same!

Observation 2. When a single BBR flow competes with a single CUBIC flow at a bottleneck, there must exist some threshold bottleneck buffer size T_{fair} such that when the bottleneck buffer size $Buff < T_{fair}$, the BBR flow gets higher throughput than the CUBIC flow and when $Buff > T_{fair}$, the CUBIC flow gets higher throughput than BBR.

Throughput in relation to number of flows

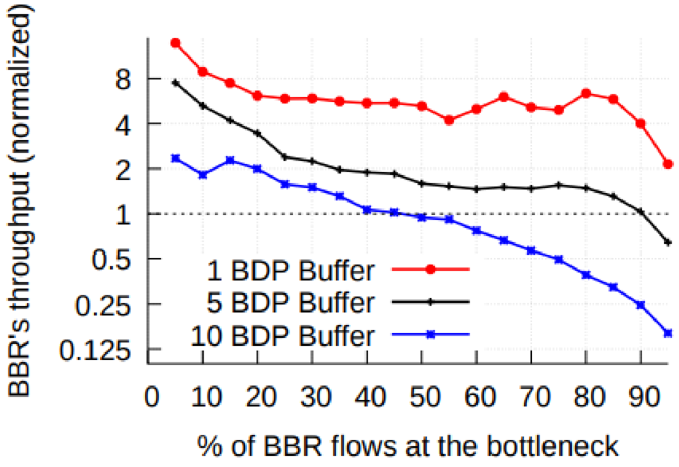


Fig. 2. BBR's throughput vs. % of BBR flows.

- Empirically, we can also observe that as the percentage of BBR flows at the bottleneck increases, their per-flow average throughput reduces.

Observation 3. As the percentage of BBR flows at the bottleneck increases, the per-flow average throughput of BBR flows at that bottleneck decreases.

Throughput in relation to RTT

- CUBIC flows with a smaller RTT are able to probe much faster than flows with longer RTTs due to frequent feedback.
- BBR flows with longer RTT are able to get higher bandwidth than flows with smaller RTTs because they place 1 BDP worth of packets in the bottleneck buffer

> **Observation 4.** When two BBR flows compete at a bottle-neck, the flow with a longer RTT will get higher bandwidth than the flow with a shorter RTT. When two CUBIC flows compete at a bottleneck, the flow with a shorter RTT will get higher bandwidth than the flow with a longer RTT.

Using these observation to predict a Nash Equilibrium in a 2 flow game

- A Nash Equilibrium in a game is a strategy distribution between the players where no player has anything to gain by changing only their strategy.

Table 2. Outcomes in a two-flow game. ($RTT_1 > RTT_2$, winning strategies are **highlighted**)

	$Buff < T_2$				$T_2 < Buff < T_3$				$T_3 < Buff$			
Scenario	Strategies		Outcome		Strategies		Outcome		Strategies		Outcome	
	S_1	S_2	S_1	S_2	S_1	S_2	S_1	S_2	S_1	S_2	S_1	S_2
1	C	**C**	L	W	C	**C**	L	W	C	**C**	L	W
2	C	**B**	L	W	**C**	B	W	L	**C**	B	W	L
3	**B**	C	W	L	**B**	C	W	L	B	**C**	L	W
4	**B**	B	W	L	**B**	B	W	L	**B**	B	W	L
Nash Equilibrium												

Nash Equilibria in a general *n*-flow game

- The complete mathematical proof is beyond the scope of this work, we therefore make the conjecture that a NE will exist in an n-flow game
- This conjecture is based on the observation that BBR flows get diminishing returns in throughput as the percentage of flows at the bottleneck increase

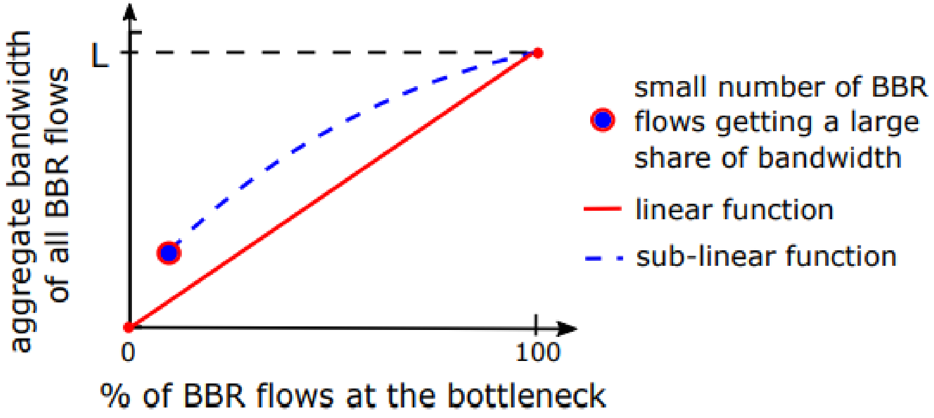


Fig. 3. Sub-linear increase in total BBR bandwidth.

Checking the claims of this conjecture in a limited state space

- We ran various number of flows through a common bottleneck link and measured their throughputs.
- For a given number of flows and a network configuration we ran all the possible combinations of flows running either CUBIC or BBR.
- A given distribution of CUBIC and BBR was considered to be the NE if in that combination, all of the flows got worse throughput if they switched their congestion control algorithm (while everyone else ran the same CC)

CBC

Is the NE if:

In BBC, flow 1 gets worse throughput
In CCC, flow 2 gets worse throughput
In CBB, flow 3 gets worse throughput

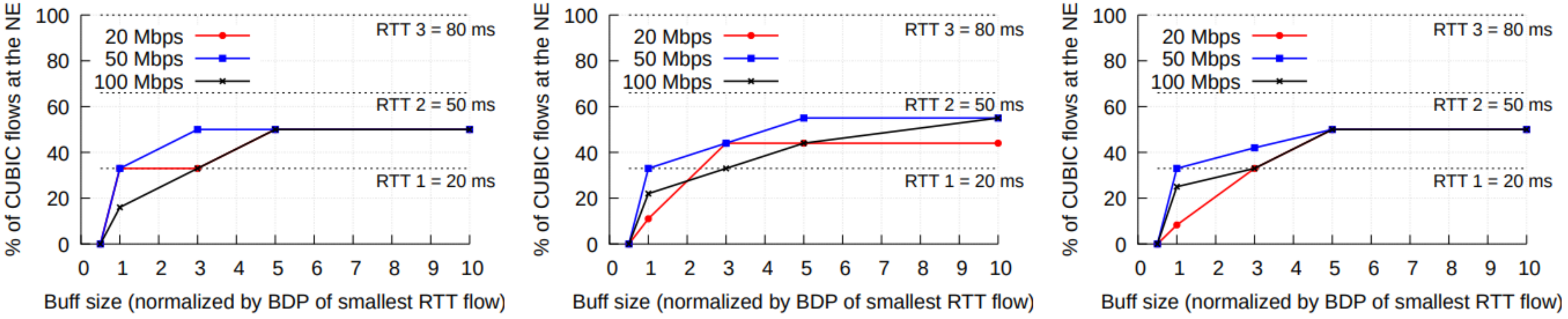
Properties of observed NE

- NE was computed in 6, 9 and 12 flow systems with each third of the flows having RTTs 20, 50 and 80 ms respectively
- In each case, we observed exactly one Nash Equilibrium
- In each Nash Equilibrium, when the flows were sorted by the RTT, CUBIC was always picked by the smallest RTT flows

(CCC…)(…BBB)

m flows (n-m) flows

Effect of buffer size and link speed on the NE

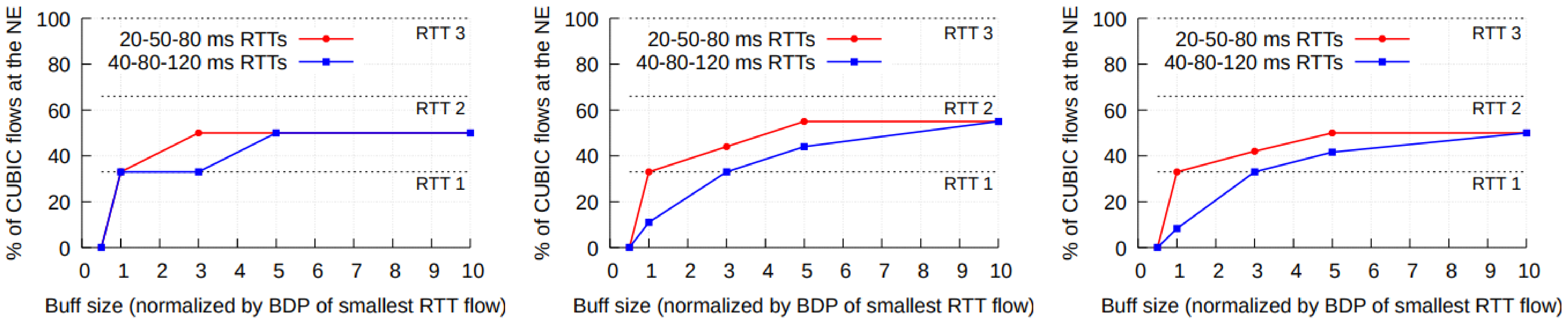


(a) 6 flows (b) 9 flows (c) 12 flows

Fig. 4. The effect of link capacity and number of flows on the Nash Equilibrium. RTTs 20 ms, 50 ms and 80 ms.

- Buffer size had the biggest effect on where the NE was
- At high buffer sizes, the NE seems to be at a 50-50 split between CUBIC and BBR

Effect of RTT distribution on the NE



(a) 6 flows, 50 Mbps (b) 9 flows, 50 Mbps (c) 12 flows, 50 Mbps

Fig. 5. The effect of the RTT distribution on the Nash Equilibrium.

- RTT distribution had a small effect of the where the NE was

Conclusion

- Despite BBR's current throughput benefits, CUBIC is unlikely to disappear soon from the Internet
- The Internet is likely to remain a heterogeneous mix of congestion control algorithms
- TCP performance is highly contextual

Future work

- Formal proof for NE is a general n-flow game
- The Internet does not follow economic game theory
- Effect on the NE in the presence of multi-hop paths, large number of flows and AQMs

Thank you!
Questions?