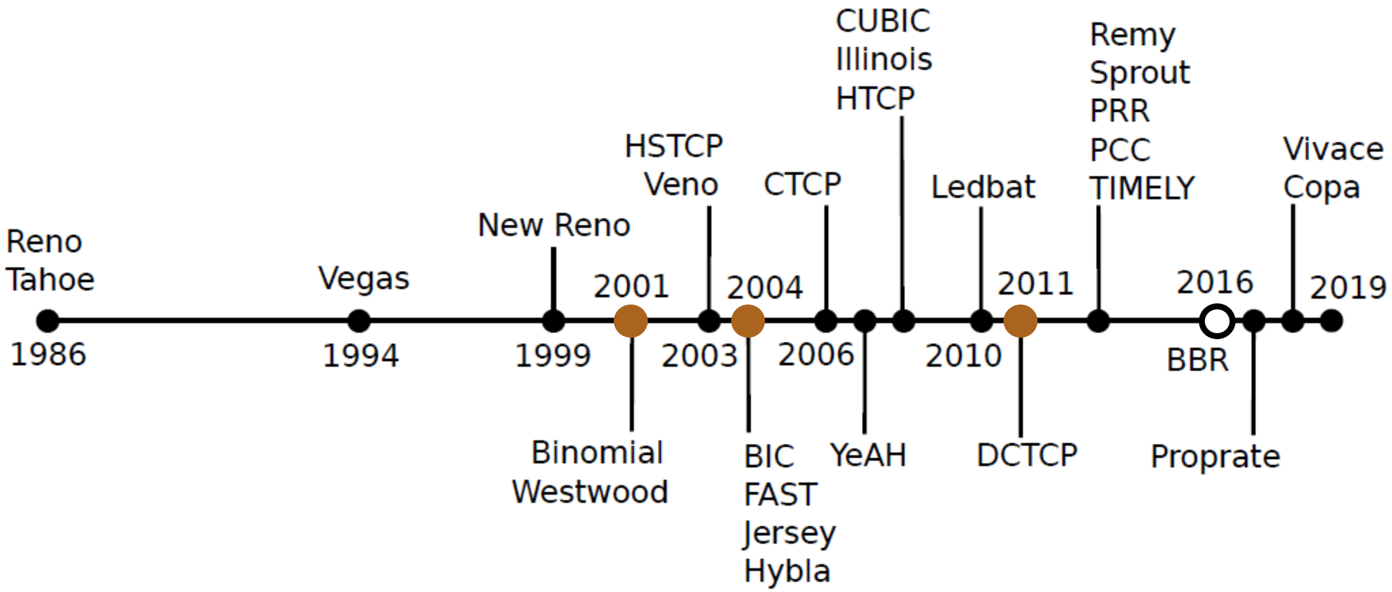
THE GREAT INTERNET TCP CONGESTION CONTROL CENSUS

Ayush Mishra

IETF 109
20th November, 2020

THIRTY YEARS OF CONGESTION CONTROL ON THE INTERNET.



● Years in which similar TCP censuses were conducted


Google Cloud

Blog Latest Stories Product News Topics

GOOGLE CLOUD PLATFORM

TCP BBR congestion control comes to GCP – your Internet just got faster

Neal Cardwell
Senior Staff Software Engineer

Yuchung Cheng
Senior Staff Software Engineer

C. Stephen Gunn
Senior Staff Site Reliability Engineer

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 often delivers at only a few Mbps over intercontinental distances.[6]

These problems result from a design choice made

AIM:
CONDUCT A CONGESTION CONTROL CENSUS AMONG THE 20,000 MOST POPULAR WEBSITES* ON THE INTERNET.

*ACCORDING TO THEIR ALEXA RANKINGS

THIS IS A NON-TRIVIAL TASK

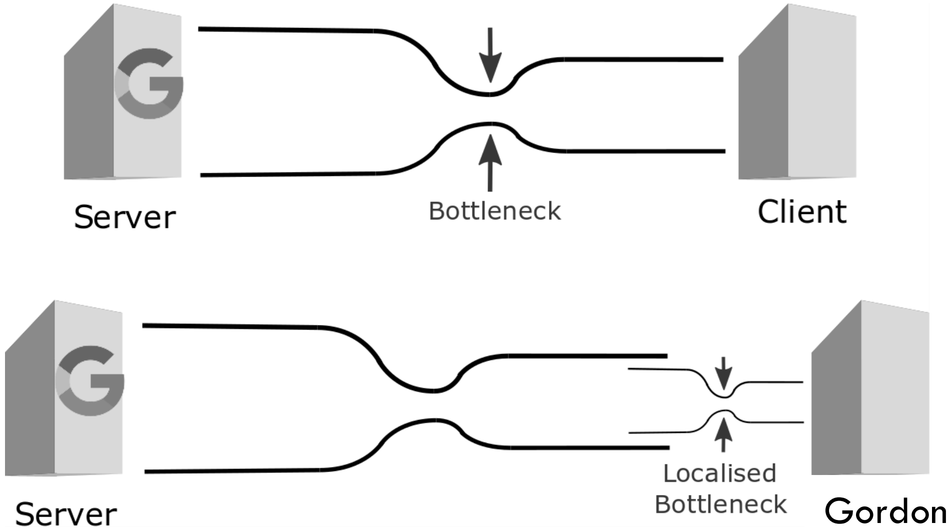
The measurement tool would need to

1. **Isolate** the Internet's **network dynamics**.
2. **Extract a common feature** from a variety of congestion control algorithm.
3. Identifying congestion control algorithm behavior within **short HTTP page downloads**.

OUR SOLUTION: GORDON

1

ISOLATING THE INTERNET'S NETWORK DYNAMICS



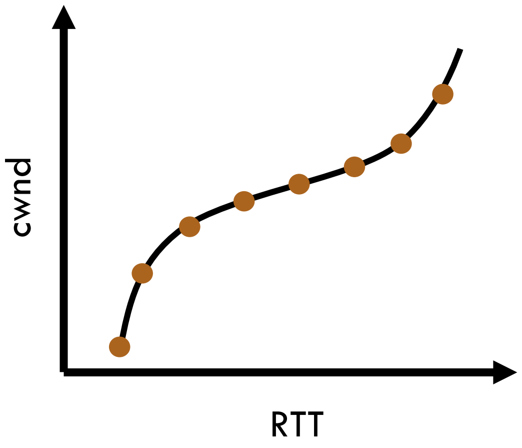
Most network dynamics, like change in bottleneck bandwidth and packet losses happen at the connection bottleneck. Localizing the bottleneck allows us to have better control over the connection.

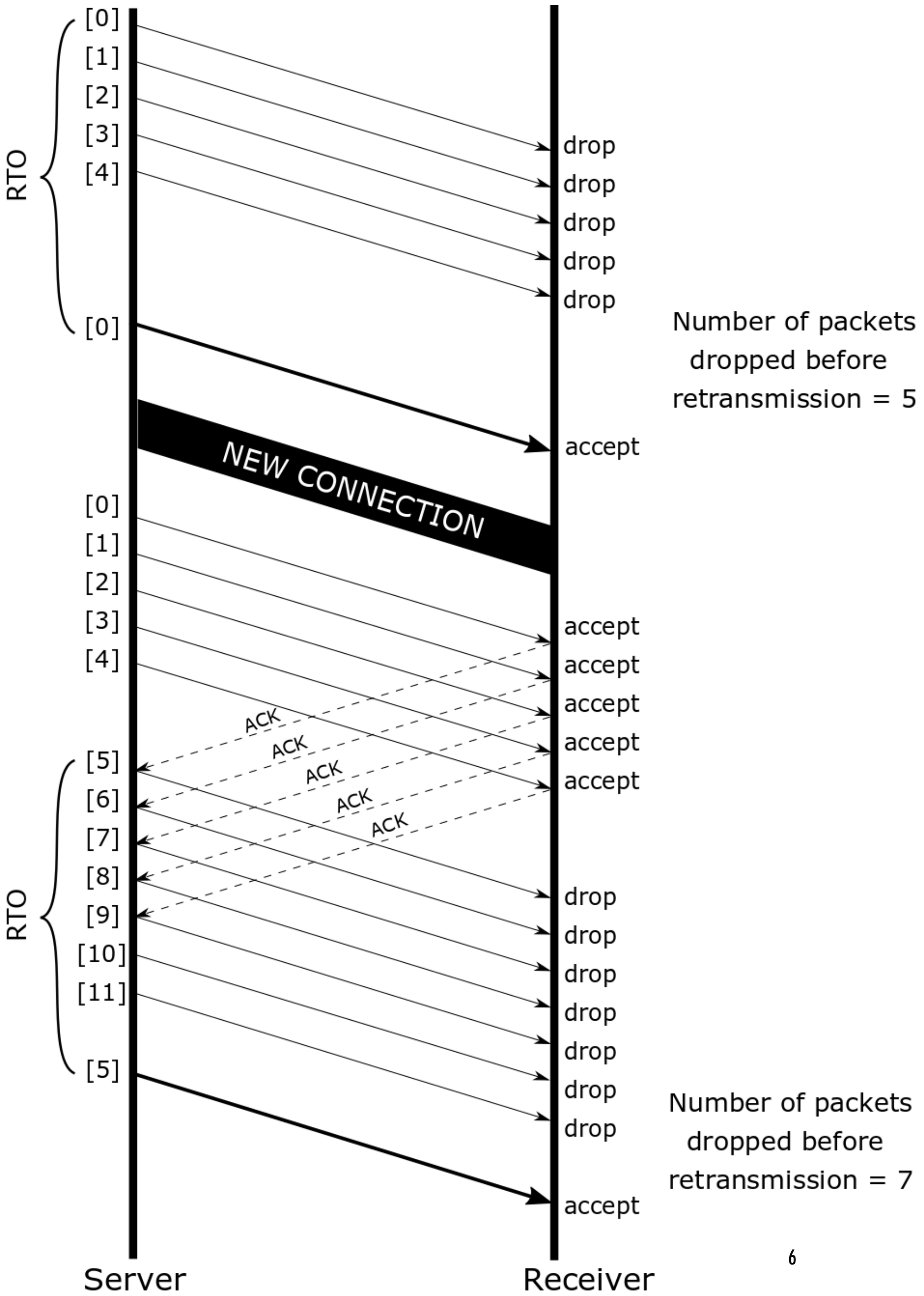
2

EXTRACTING A COMMON FEATURE

Gordon captures a remote sender's **cwnd evolution graph**.

It does so by **dropping packets** to estimate the maximum tolerable in-flight packets (cwnd).

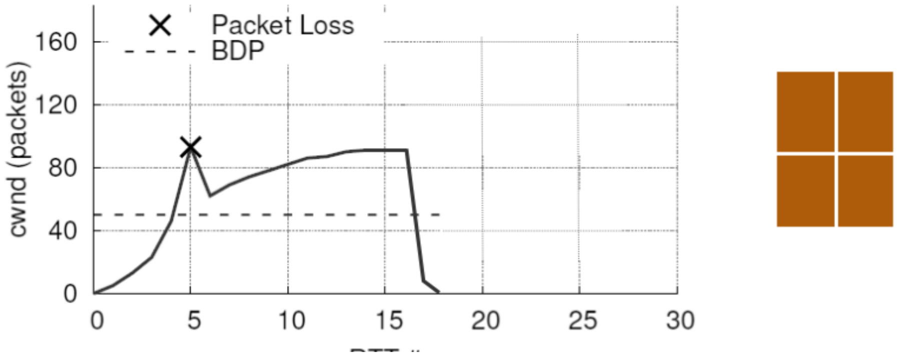


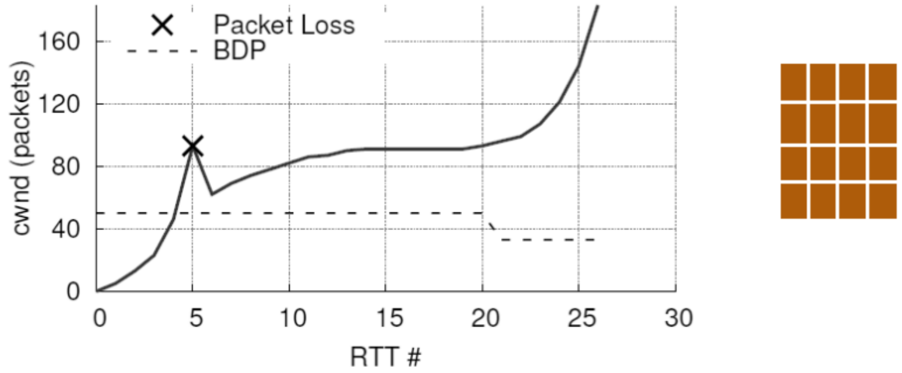


3

DEALING WITH
SHORT HTTP PAGE DOWNLOADS

1. We crawled the target domains for the **largest pages** we could find
2. We used the **smallest MTU** allowed by the network path



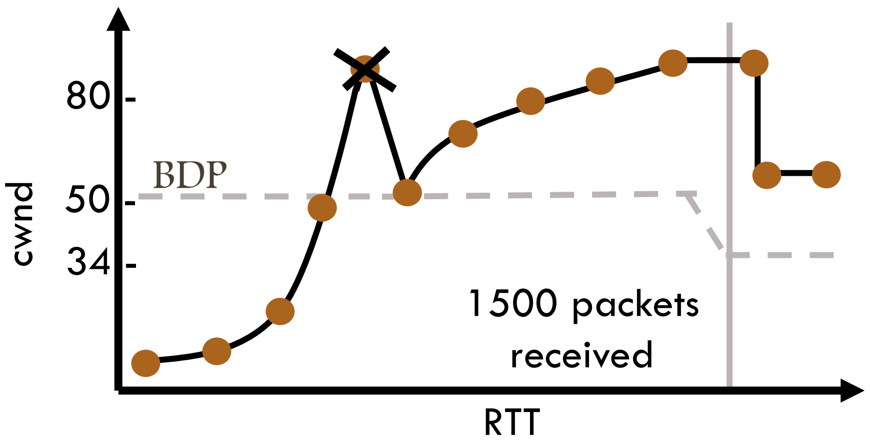


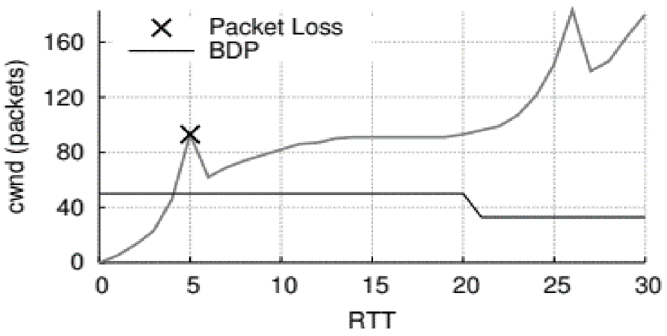
SIMULATING NETWORK STIMULI

Gordon emulates **2 key network stimuli** to elicit characteristic responses from congestion control algorithms.

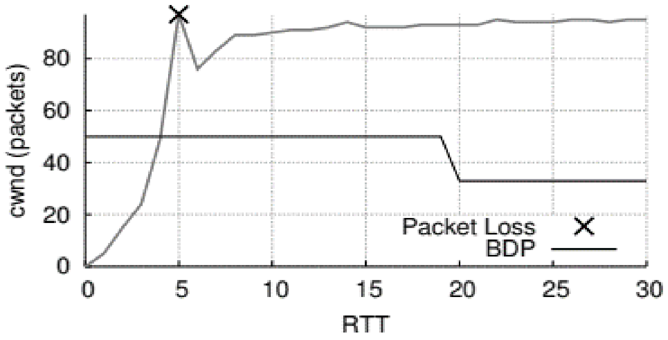
We call these set of network stimuli the **Network Profile**

1. **Packet drop** at the first cwnd that exceeds 80 packets
2. **Bandwidth change** after receiving 1500 packets
3. Emulating an RTT of 100 ms

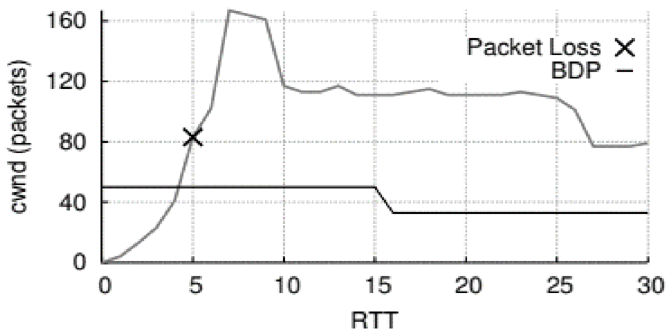




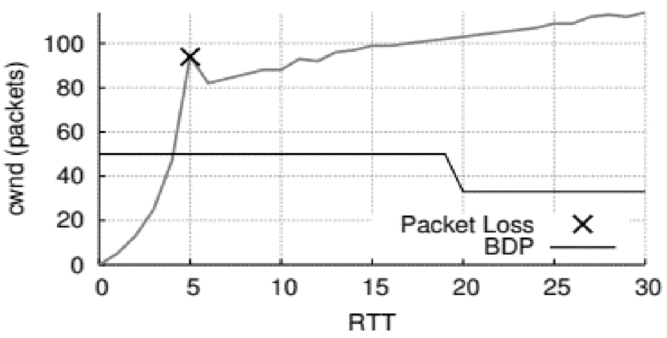
(a) CUBIC



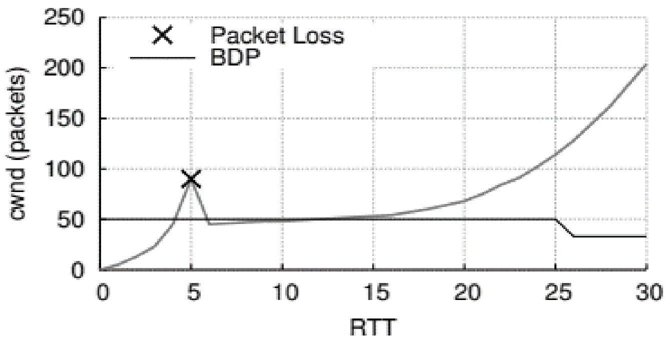
(b) BIC



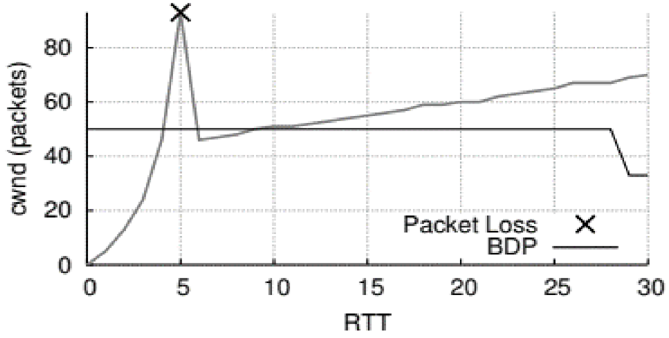
(c) BBR



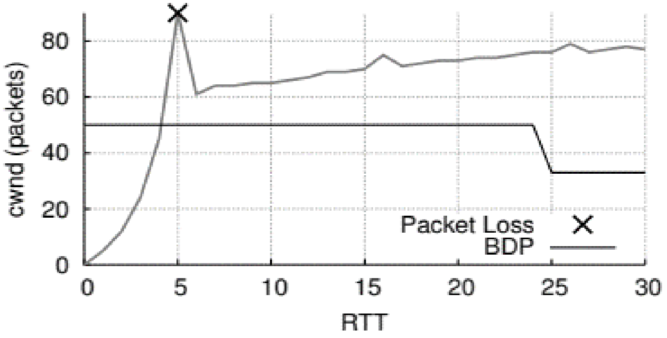
(d) Scalable



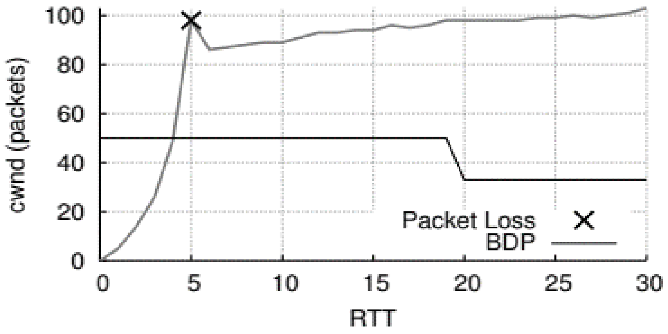
(e) HTCP



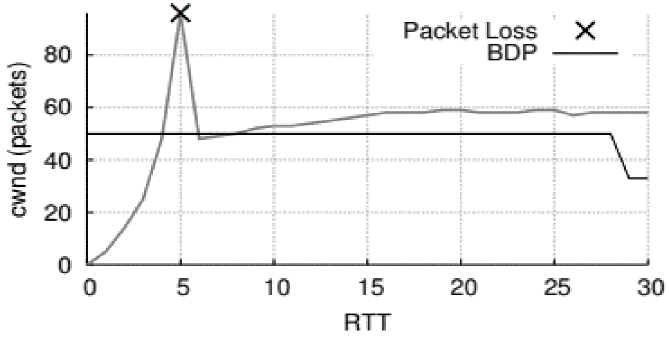
(f) New Reno



(g) Illinois



(h) YeAH



(i) Vegas

MEASUREMENT RESULTS

MEASUREMENT ACCURACY

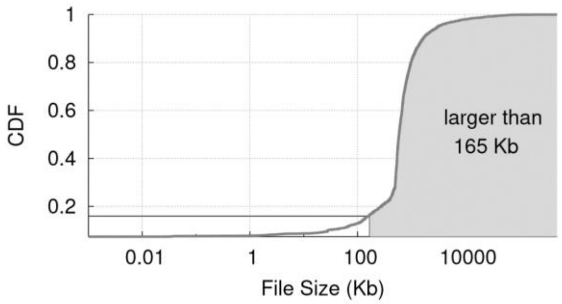
	Classified as								
	BBR	CUBIC	BIC	HTCP	Scalable	YeAH	Vegas	New Reno/Veno	CTCP/Illinois	Unknown
BBR	**98%**	0%	0%	0%	0%	0%	0%	0%	0%	2%
CUBIC	0%	**95%**	0%	0%	0%	0%	0%	0%	0%	5%
BIC	0%	9%	**91%**	0%	0%	0%	0%	0%	0%	0%
HTCP	0%	0%	0%	**95%**	0%	0%	0%	0%	0%	5%
Scalable	0%	0%	0%	0%	**98%**	0%	0%	0%	0%	2%
YeAH	0%	0%	2%	0%	0%	**98%**	0%	0%	0%	0%
Vegas	0%	0%	0%	0%	0%	0%	**94%**	6%	0%	0%
New Reno/Veno	0%	0%	0%	0%	0%	0%	0%	**96%**	0%	4%
CTCP/Illinois	0%	0%	3%	0%	0%	0%	0%	0%	**94%**	3%

MEASUREMENT **STATISTICS**

The measurements were done between **11 July 2019** and **17 October 2019** from servers in **Singapore**, **Mumbai**, **Paris**, **Sao Paulo** and **Ohio**.

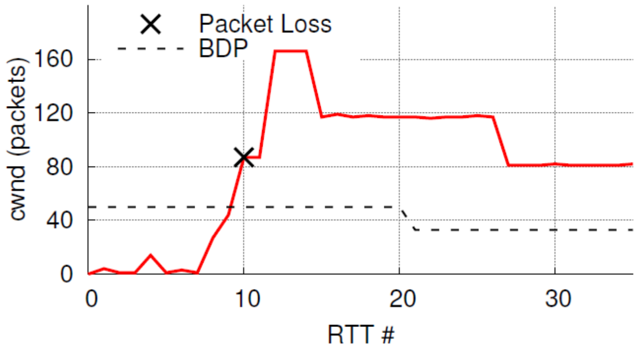
Given our network profile, 16% of pages were less than optimal in size of 165 Kb (**Short flows**)

We also came across 1,302 **Unresponsive websites**

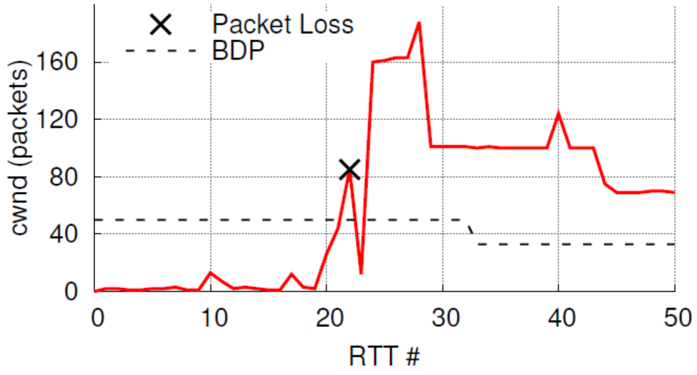


DISTRIBUTION BY WEBSITE COUNT

Variant	Websites	Proportion
CUBIC [15]	6,139	30.70%
BBR [4]	3,550	17.75%
BBR G1.1	167	0.84%
YeAH [2]	1,162	5.81%
CTCP [34]/Illinois[22]	1,148	5.74%
Vegas [3]/Veno [13]	564	2.82%
HTCP [21]	560	2.80%
BIC [37]	181	0.90%
New Reno [28]/HSTCP [12]	160	0.80%
Scalable [20]	39	0.20%
Westwood [7]	0	0.00%
Unknown	3,535	17.67%
Short flows	1,493	7.46%
Unresponsive websites	1,302	6.51%
Total	20,000	100%

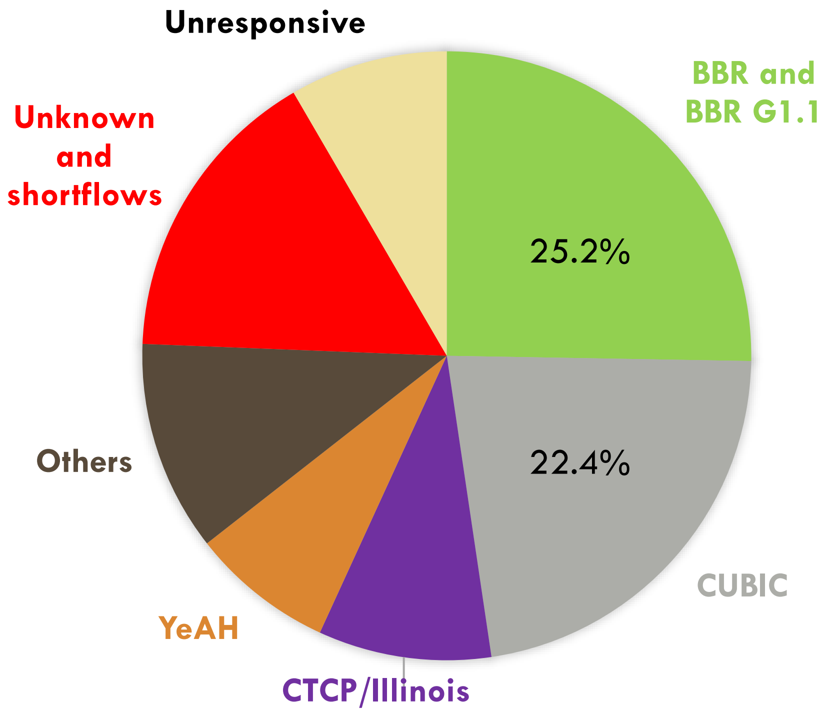


BBR



BBR G1.1

DISTRIBUTION BY POPULARITY AND TRAFFIC SHARE



Share of congestion control algorithms deployed by website count in the **Alexa Top 250** websites

- Among the top 250 Alexa websites, BBR has a larger share by website count than Cubic
- In terms of traffic share, BBR is now contributing to **more than 40%** of the downstream traffic on the Internet!

Site	Downstream traffic share	Variant
Amazon Prime	3.69%	CUBIC
Netflix	15%	CUBIC
YouTube	11.35%	BBR
Other Google sites	28%	BBR
Steam downloads	2.84%	BBR

(As measured on **static HTTP webpages**)

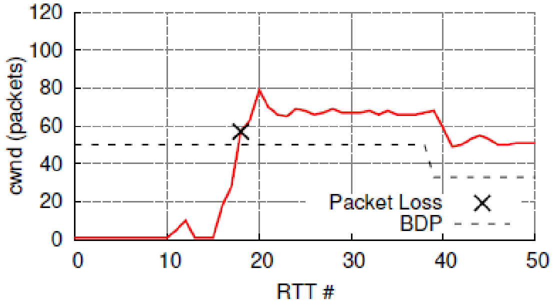
LOOKING CLOSER AT THE UNCLASSIFIED VARIANTS

We had a total of **6,330 (31.65%)** of websites that were **unclassified**

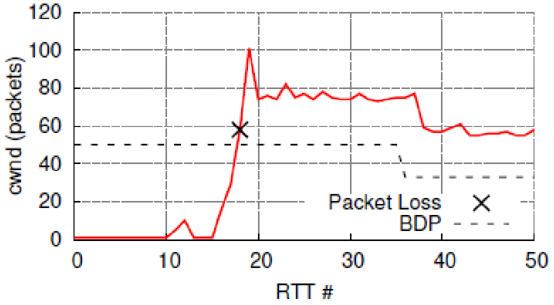
We ran a variety of network profiles on these websites to infer something about their congestion control mechanism

Type	React to Packet Loss?	React to BDP?	Websites (share)
AkamaiCC	✗	✓	1,103 (5.52%)
Unknown Akamai	✗	?	157 (0.79%)
Unknown	✗	?	493 (2.47%)
	✓	?	1,782 (8.91%)
Short flows	✓	?	1,493 (7.47%)
Unresponsive	?	?	1,302 (6.51%)
Total			6,330 (31.65%)

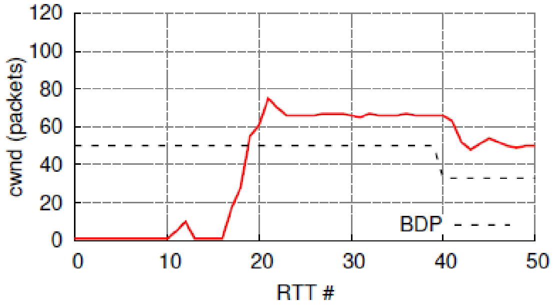
AKAMAI CC AND OTHER DOMAIN SPECIFIC VARIANTS



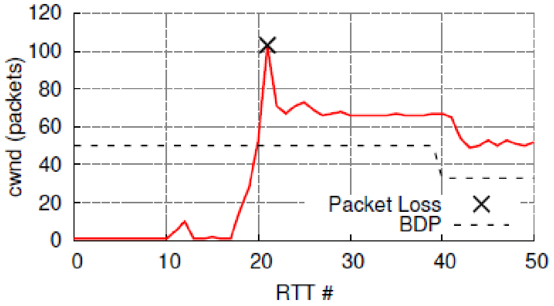
(a) Custom Network Profile 4 - Shape 1.



(b) Custom Network Profile 4 - Shape 2.

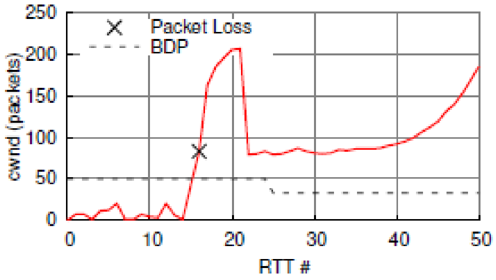


(c) Custom Network Profile 1 - Shape 1.

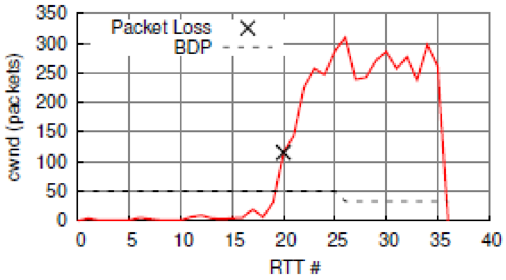


(d) Custom Network Profile 1 - Shape 2.

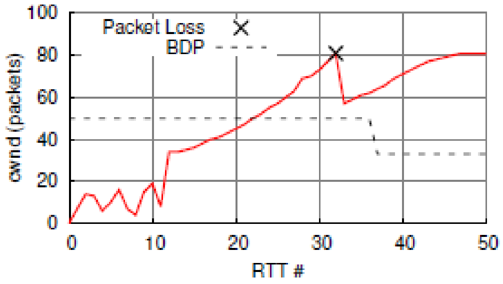
Akamai CC



(a) amazon.com

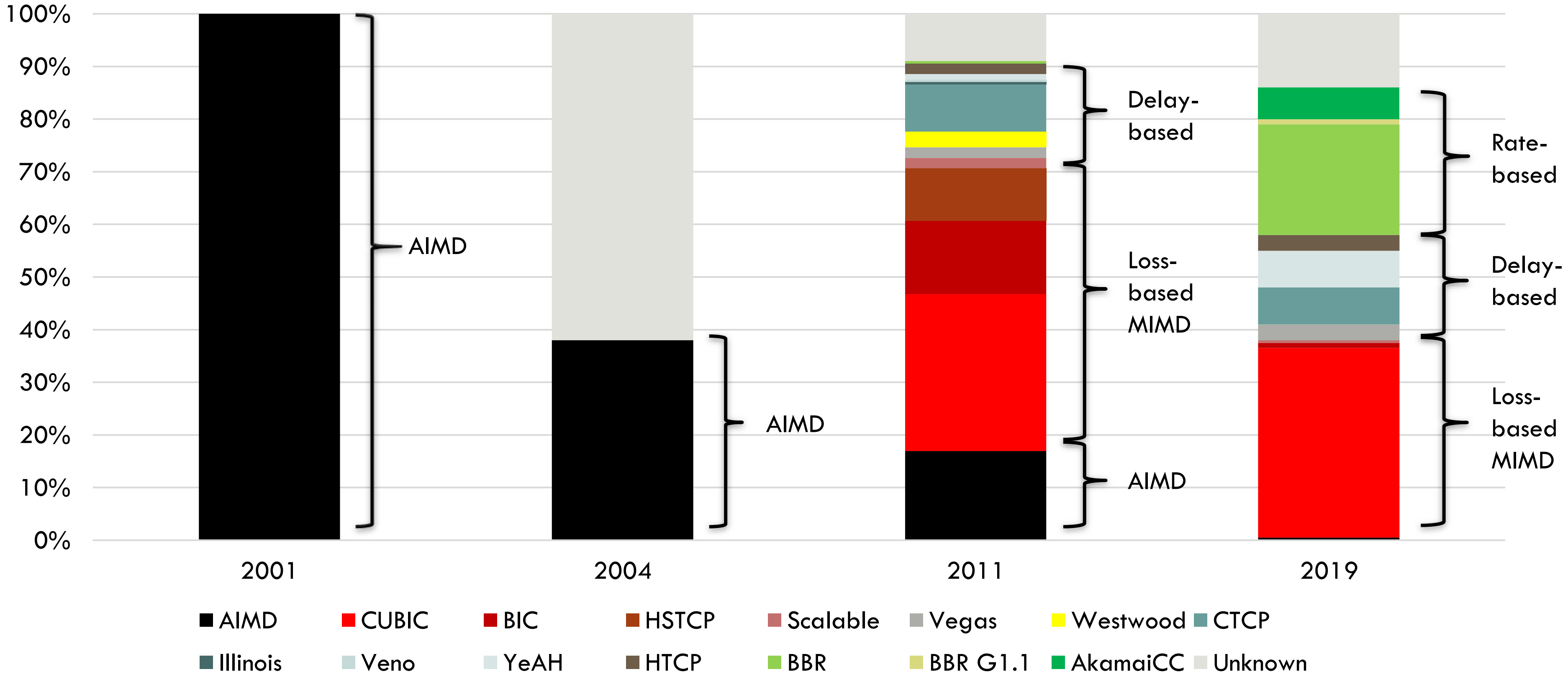


(b) zhihu.com



(c) yahoo.co.jp

THE EVOLUTION OF THE TCP ECOSYSTEM



FUTURE WORK ON GORDON

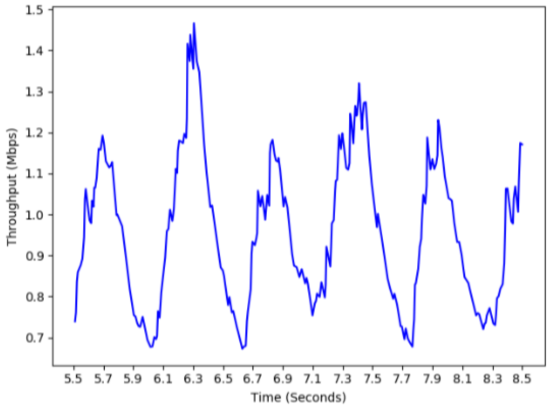
Investigate Unresponsive websites to increase data set

Experiment with **other Network stimuli**

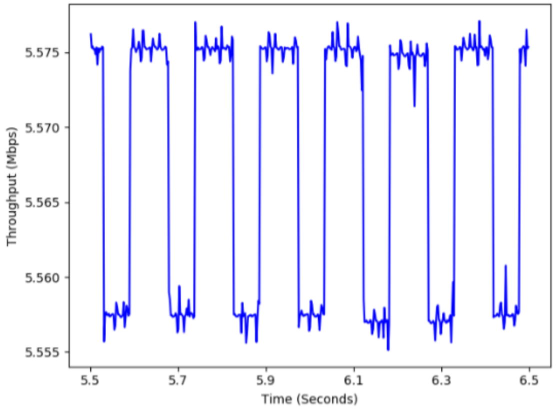
Extending the tool to **detect sub-RTT behaviors**

Using clustering algorithms to **better analyze the Unknown websites**

Identifying **other rate-based** algorithms, beyond BBR



Copa

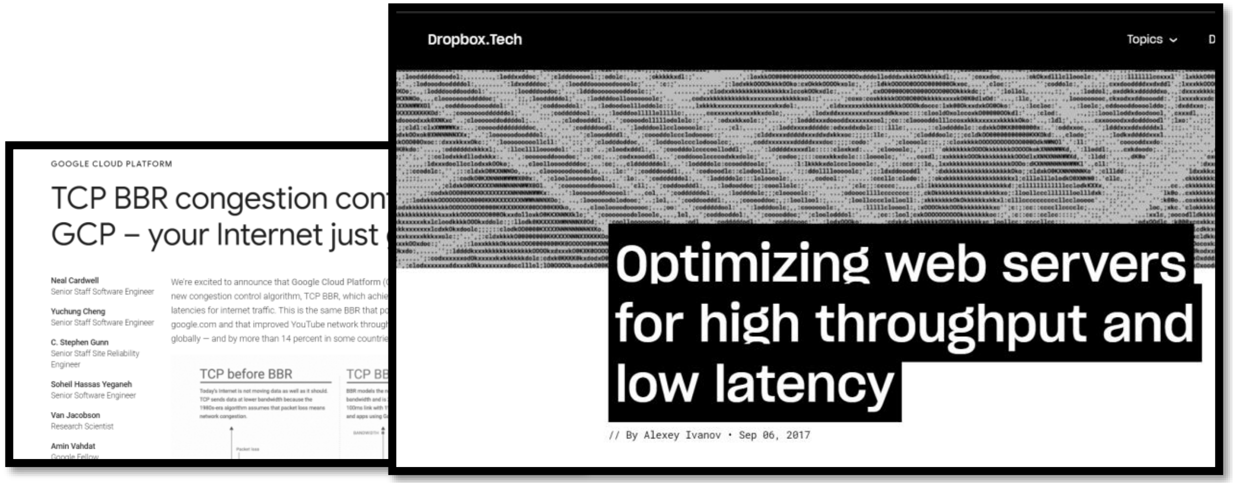


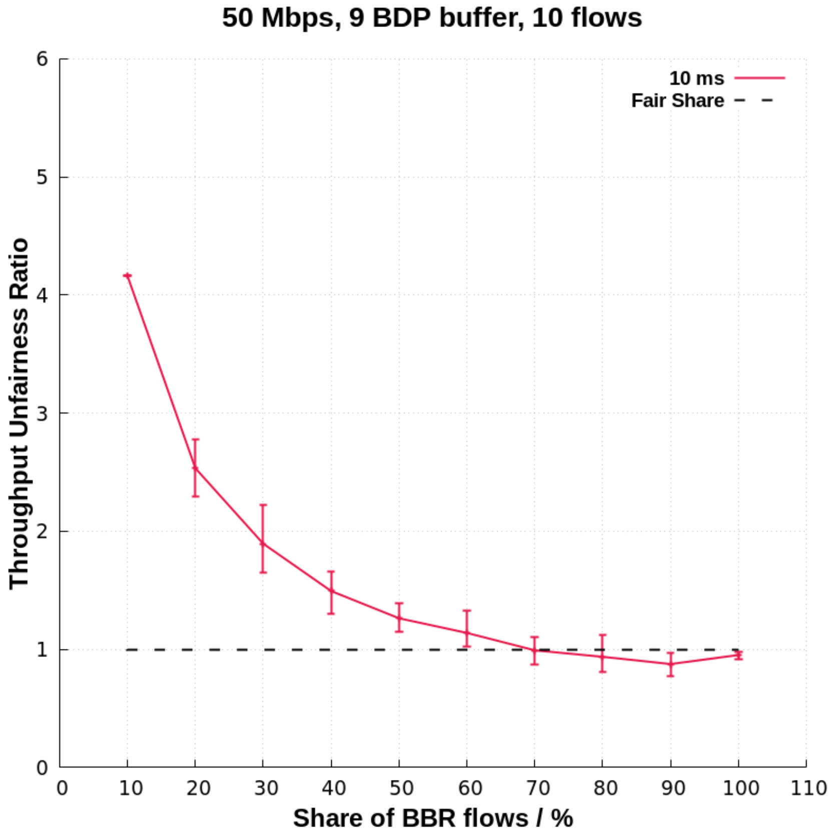
PCC-Vivace

RESEARCH QUESTION: HOW WILL BBR AND CUBIC FAIR IN THIS EVOLVING CC LANDSCAPE?

There has been plenty of work that shows how BBR can be unfair to CUBIC in certain network scenarios, but how will an evolving CC landscape effect BBR?

Will BBR's performance benefits sustain?





RESEARCH QUESTION: UNDERSTANDING THE RATE-BASED CC MECHANIC

BBR and **other newer Internet Congestion Control algorithms** proposed since are all **rate-based**.

All these algorithms work on tight **send-rate** and **receive-rate feedback loops**.

Classic congestion control questions on **fairness guarantees** and **convergence** need to be re-answered for the new rate-based congestion control paradigm.

THANK YOU

I look forward to your questions!

Please email me on
ayush@comp.nus.edu.sg