Buffer sizing and Video QoE Measurements at Netflix

Bruce Spang, Brady Walsh, Te-Yuan Huang, Tom Rusnock, Joe Lawrence, Nick McKeown

February 10, 2020
What are we talking about?
What are we talking about?
How big should a buffer be?

**Too big:** packets wait for too long

**Too small:** too many packets thrown away
“A buffer should be at least one BDP” [Villamizar, Song 1994]
“A buffer should be at least one BDP” [Villamizar, Song 1994]

BDP = Bandwidth x Delay
# of packets in a link for full utilization
“A buffer should be at least one BDP” [Villamizar, Song 1994]

BDP = Bandwidth x Delay
# of packets in a link for full utilization

Congestion Window

Time
“A buffer should be at least one BDP” [Villamizar, Song 1994]

BDP = Bandwidth x Delay
# of packets in a link for full utilization

Loss happens when link and buffer are full
“A buffer should be at least one BDP” [Villamizar, Song 1994]

BDP = Bandwidth x Delay

# of packets in a link for full utilization

Loss happens when link and buffer are full

TCP stops sending until $\frac{1}{2} (BDP + B)$ packets received
“A buffer should be at least one BDP” [Villamizar, Song 1994]

BDP = Bandwidth x Delay
# of packets in a link for full utilization

Loss happens when link and buffer are full

Buffer needs to hold this many packets

TCP stops sending until \( \frac{1}{2} (BDP + B) \) packets received

\[ \text{Time} \]

\[ \frac{1}{2} (BDP + B) \]

BDP + B
How big should a buffer be?

**BDP:** Villamizar and Song 1994

**BDP/√n:** Appenzeller, McKeown, Keslasy 2004

**O(n):** Dhamdhere, Jiang, Dovrolis 2005

**O(1):** Enachescu, Ganjali, Goel, McKeown, Roughgarden 2006
Which is correct?
It’s complicated
1. TCP New Reno (mostly) behaves as expected
2. Video performance varies
3. Real routers complicate this story
Our Experiment
Catalog servers
Uses spinning disks, cheaply stores entire catalog
Offload servers

Use SSDs to serve top ~30% of content faster
These three racks are called a stack.
Make this buffer small...

...and this one large
1. TCP New Reno (mostly) behaves as expected
2. Video performance varies
3. Real routers complicate this story
Large buffer has higher latency during congested hour
Sometimes the large buffer has much higher latency
Large buffer has lower loss during congested hour

% Retrans. Bytes

Buffer Size (MB)
- 50.0
- 250.0
- 500.0
- 750.0

Normalized Load
1. TCP New Reno (mostly) behaves as expected
2. Video performance varies
3. Real routers complicate this story
Good buffer size:
+ Fewer rebuffers
+ Better video quality
+ Videos start faster

Bad buffer size:
- More rebuffers
- Worse video quality
- Videos start slower
Good buffer size:
+ Fewer rebuffers
+ Better video quality
+ Videos start faster

Bad buffer size:
- More rebuffers
- Worse video quality
- Videos start slower

This happens when buffer is too large or too small.
Site #2: A smaller buffer is better

Reducing the buffer from **500MB** to **25MB**
- **15.6%** decrease in sessions with a rebuffer
- **5.3%** decrease in low quality video
- **13.5%** decrease in play delay
Site #3: A smaller buffer is better

Reducing the buffer from 500MB to 50MB
-22.1% decrease in sessions with a rebuffer
-7.0% decrease in low quality video
-14.8% decrease in play delay
Site #1: A smaller buffer is worse

Reducing the buffer from **500MB** to **50MB**

+46.3% increase in sessions with a rebuffer

+5.7% increase in low quality video

-5.9% decrease in play delay
1. TCP New Reno (mostly) behaves as expected
2. Video performance varies
3. Real routers complicate this story
Large buffer has higher latency during congested hour
Remember how the large buffer has much higher latency…
Servers have different very latency distributions
What I imagined
What I imagined

LIES!
Buffer architecture

Server #1
Server #2
Server #3

“Offload” VOQ

“Catalog” VOQ

2/3
100Gbps
1/3
ISP
Traffic is fairly split when load is equal

40 Gbps

“Offload” VOQ

67 Gbps

33 Gbps

100 Gbps

ISP

40 Gbps

“Catalog” VOQ

40 Gbps
When one VOQ offers less than its fair share, it sees no congestion

“Offload” VOQ

50 Gbps

50 Gbps

“Catalog” VOQ

10 Gbps

No delay!

ISP

100 Gbps

10 Gbps

90 Gbps
VOQs explain the RTT differences

- This VOQ is served faster
- This VOQ is served slower
- This VOQ is all over the place
Switches prioritize long-tail content
Switches prioritize long-tail content

Same latency during uncongested hours
Switches prioritize long-tail content

Same latency during uncongested hours

Popular content is congested

Long-tail content not congested
New scheduling algorithm!

“Offload” VOQ
Server #1
Server #2

“Catalog” VOQ
Server #3

Load-dependent
ISP
100Gbps

Load-dependent
New scheduling algorithm is more consistent
1. TCP New Reno (mostly) behaves as expected
2. Video performance varies
3. Real routers complicate this story
How big should a buffer be?
Thanks!

For more details, please see: