



Aalto University
School of Electrical
Engineering

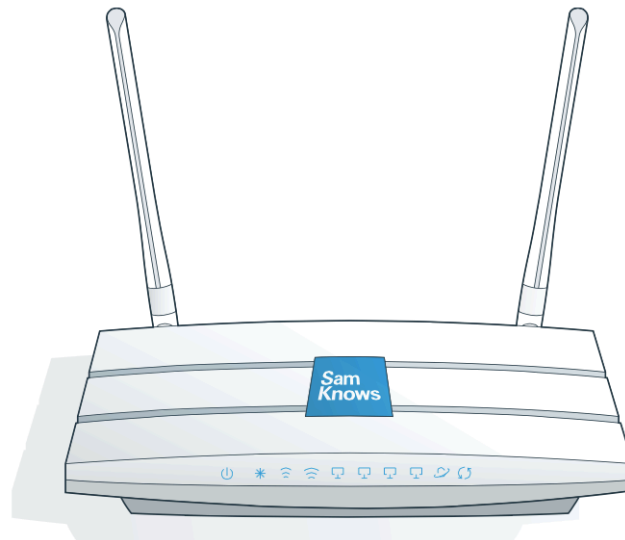
Impact of Duration on Active Video testing

Saba Ahsan

Department of Communications and Networking

Aalto University

Varun Singh (callstat.io), Jörg Ott (Technische Universität München)

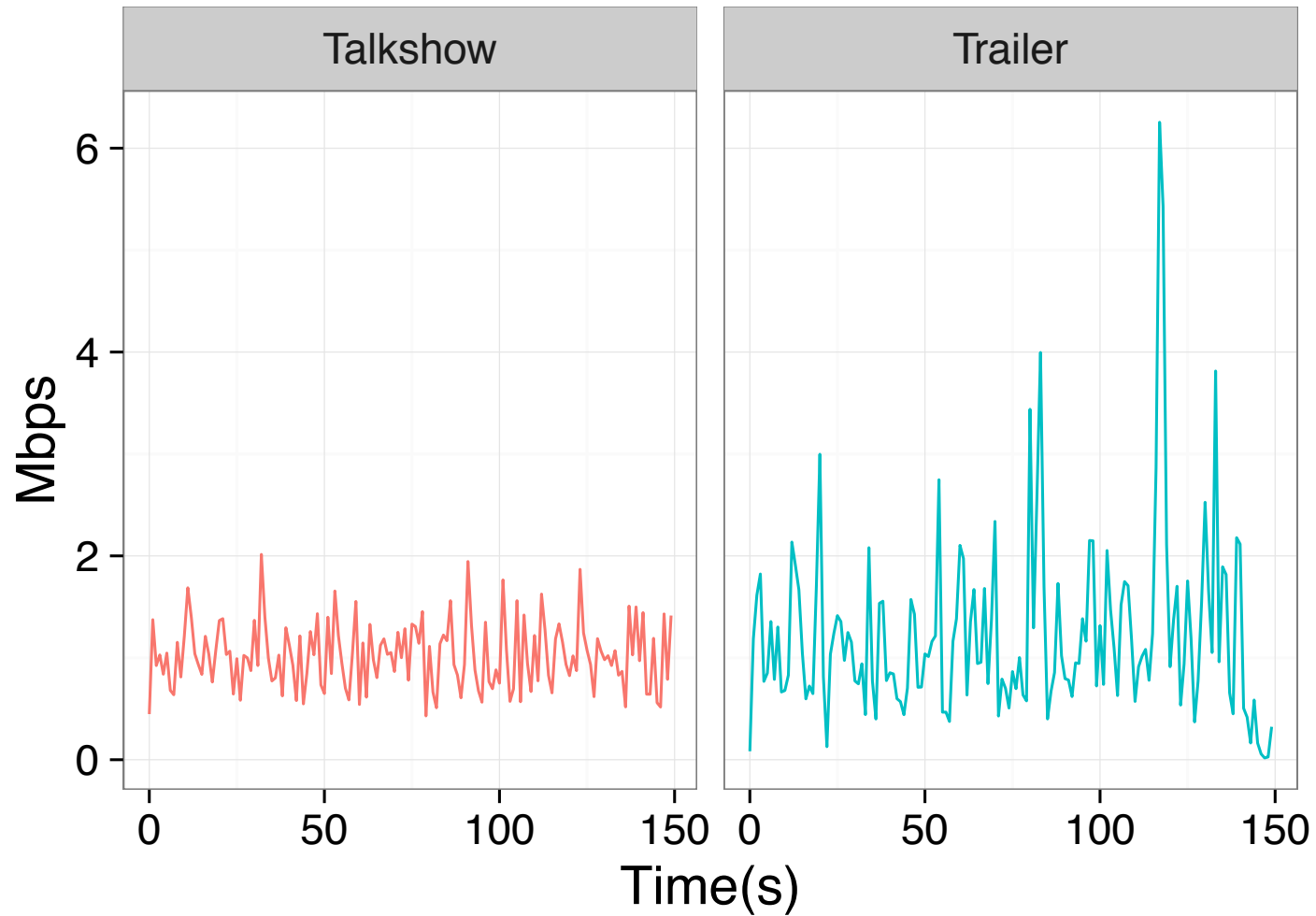


Motivation

- Actively measure performance of Internet video (YouTube) from a user's perspective
- Finish the test as quickly as possible but still get reliable results



Instantaneous video bit rate of Full HD (1080p) YouTube videos



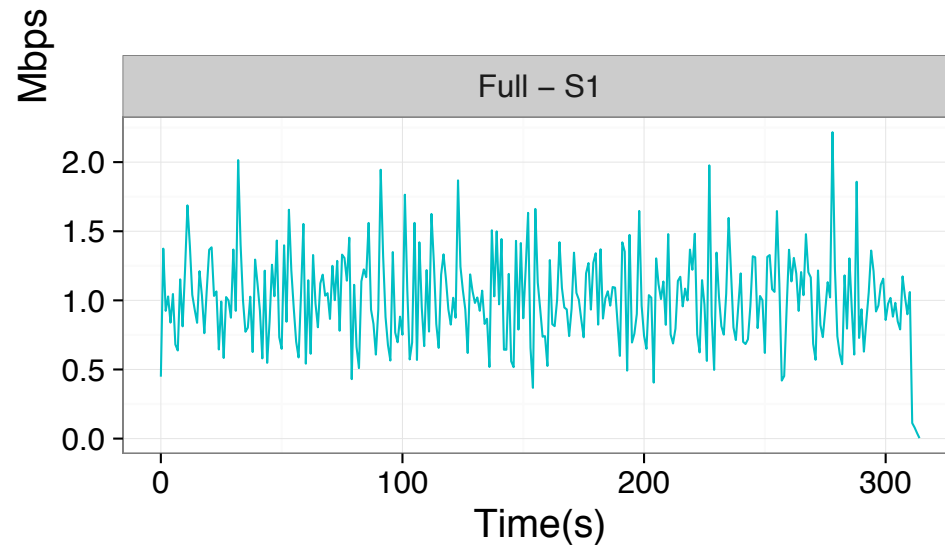
Problem Statement

- When conducting active video tests in a time constrained environment, what is the minimum that the test should run for...
 - Can we represent the bit rate magnitude and variations in Internet video with a short clip?

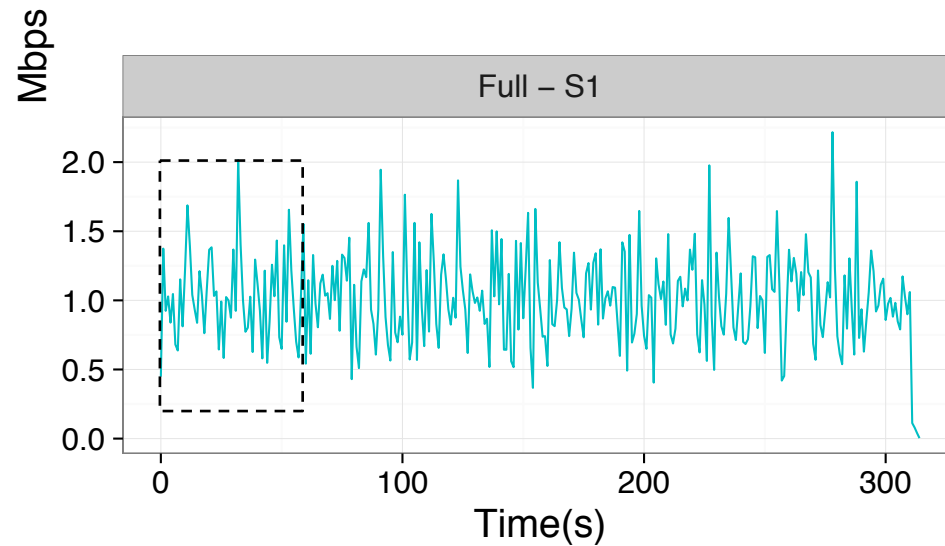
Dataset

- Popular YouTube videos by geographical area
- Mp4 videos (frame sizes and timing)
 - Non-adaptive 720p > 32k
 - Non-adaptive 360p > 62k
 - DASH 1080p > 8k
 - DASH 720p > 19k

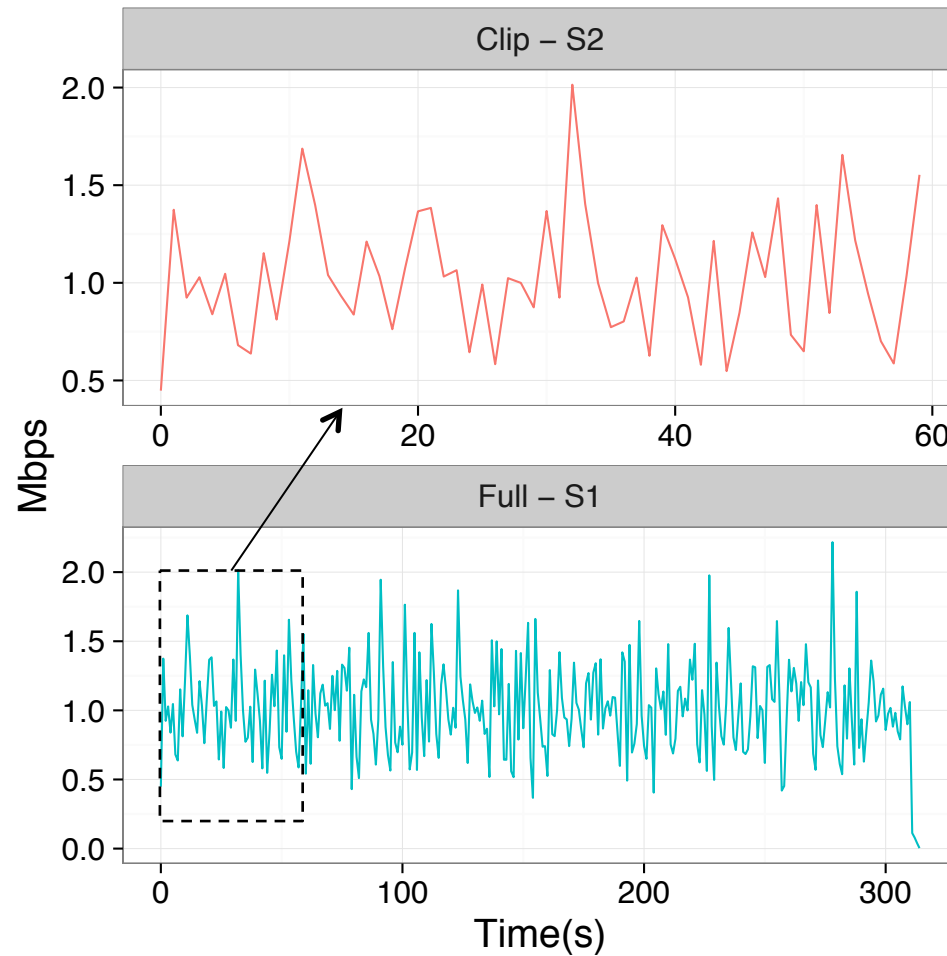
Selecting a cut-off value



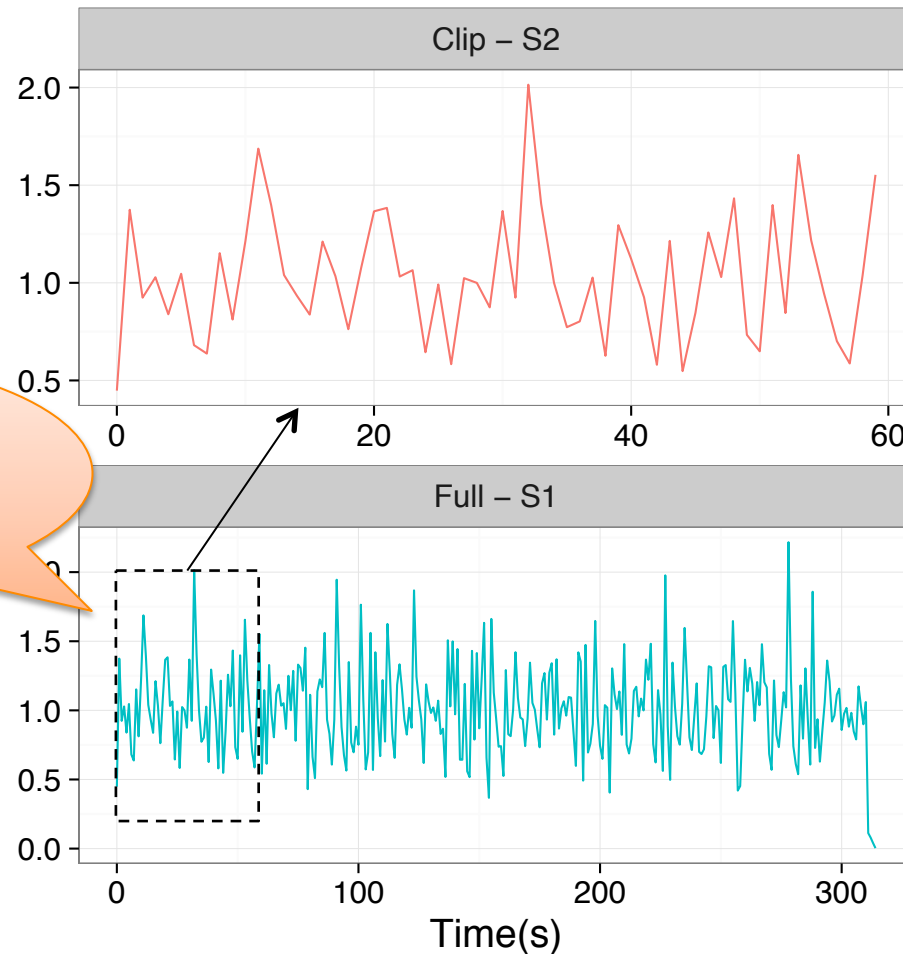
Selecting a cut-off value



Selecting a cut-off value



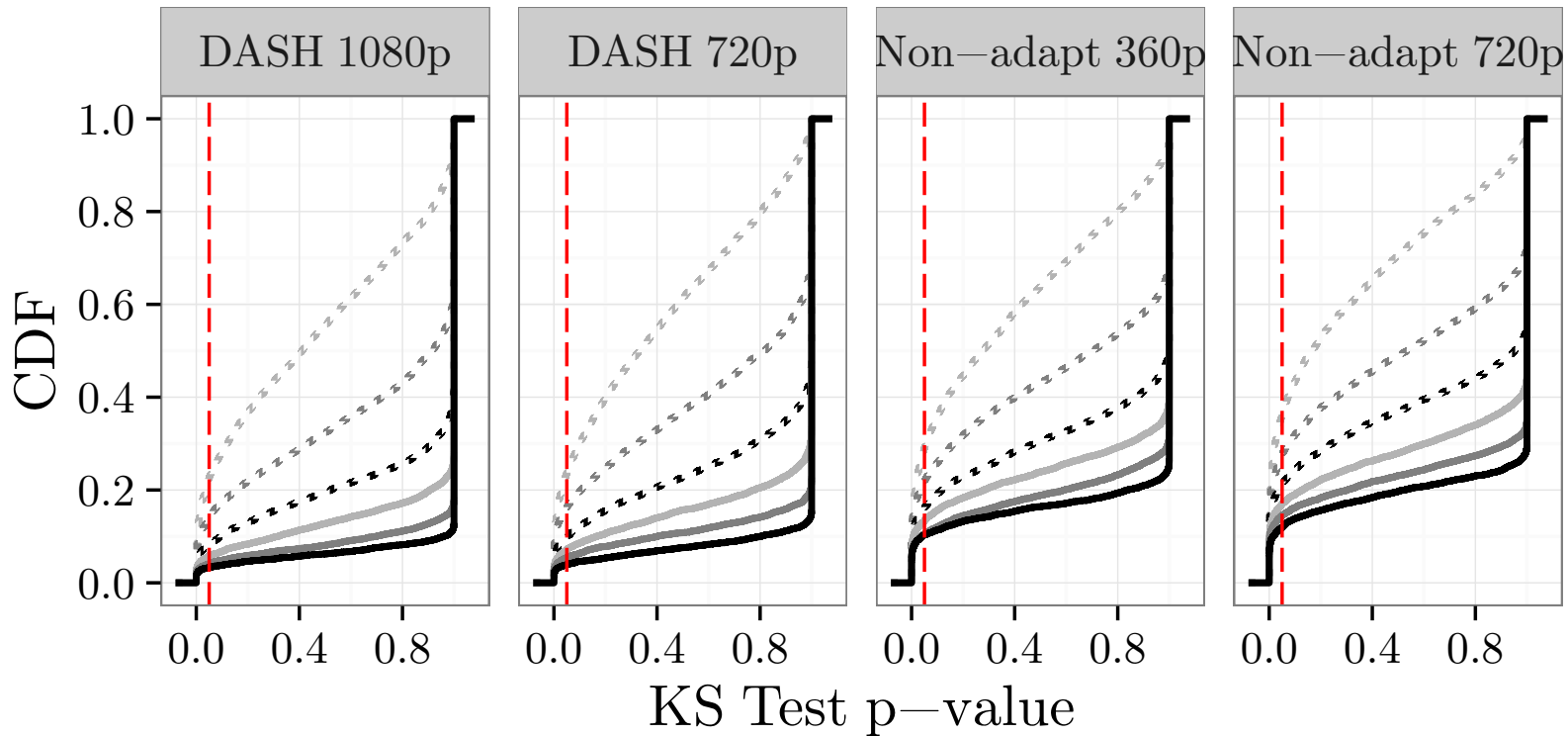
Selecting a cut-off value



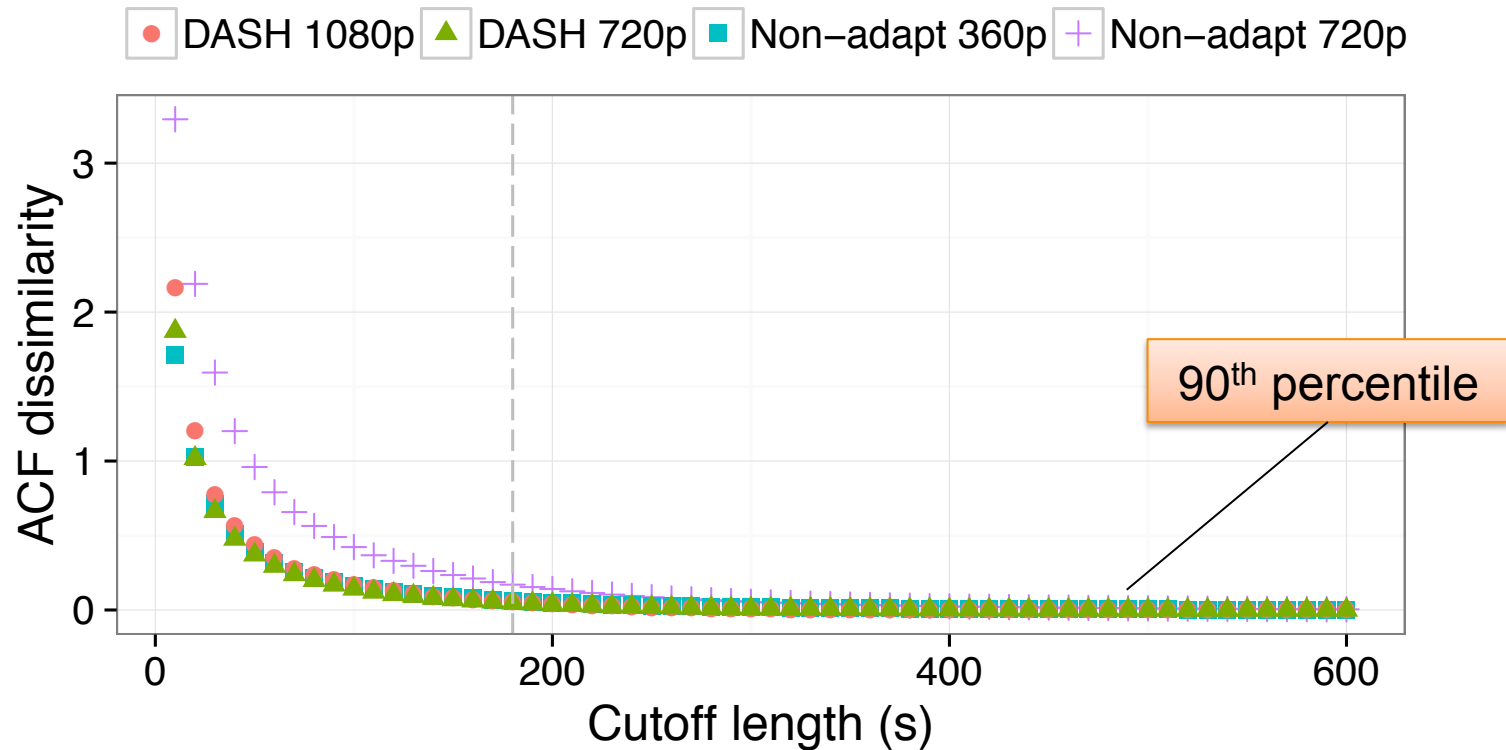
Take from the beginning, and startup delay is unchanged

Kolmogorov-Smirnov Test

Cutoff (s) · · 10 · · · 60 · · · · 120 — 180 — 240 — 300



Autocorrelation Function-based dissimilarity



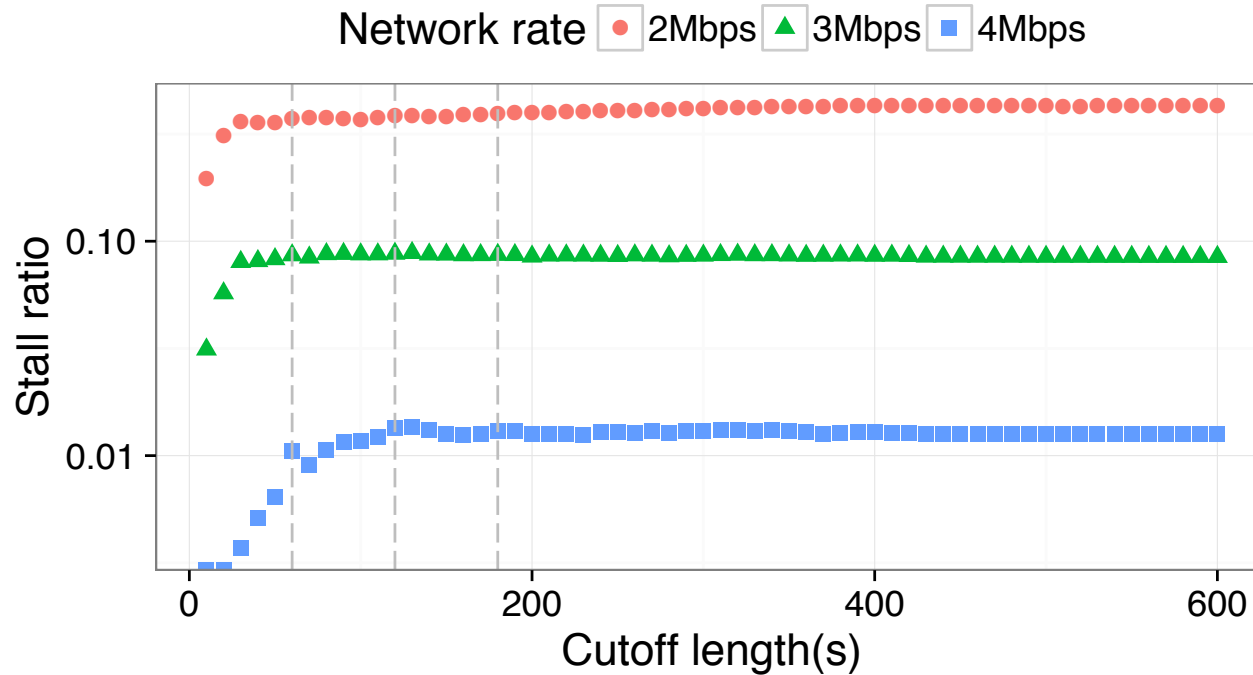
Does it really work?

- Do we get similar user experience metrics even if we cut-off the test before the video ends?
 - YES

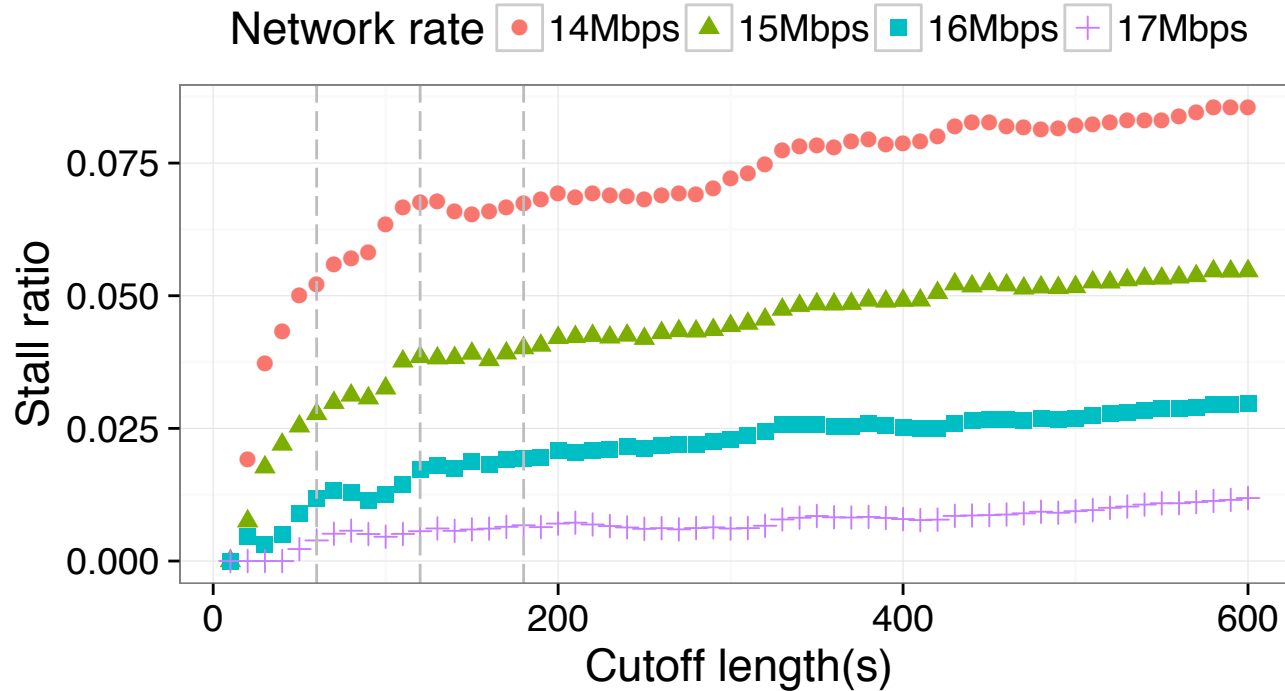
Does it really work?

- Do we get similar user experience metrics even if we cut-off the test before the video ends?
 - YES (We tested with our own client in emulated networks)
we measured stall ratio= total stall duration/total playable video

Testing in Experiment network with 10 YouTube videos of different lengths



Testing in Experiment network with 10 YouTube videos longer than 10 min (600 s)



Summary and Future work...

- Test durations should run at least for 1 min, but 3 min gives more reliable results.

Summary and Future work...

- Test durations should run at least for 1 min, but 3 min gives more reliable results.
- What about DASH related metrics?

Summary and Future work...

- Test durations should run at least for 1 min, but 3 min gives more reliable results.
- What about DASH related metrics?
- Results “probably” not suitable for feature-length movies



Aalto University
School of Electrical
Engineering

Thank you!