

Presented at LANMAN 2015



Delivery of Adaptive Bit Rate Video: Balancing Fairness, Efficiency and Quality

Jason J. Quinlan
Dept. of Computer Science,
University College Cork,
Ireland
j.quinlan@cs.ucc.ie

Ahmed H. Zahran
Dept. of Computer Science,
University College Cork,
Ireland
a.zahran@cs.ucc.ie

K. K. Ramakrishnan
Dept. of Computer Science,
University of California,
Riverside
kk@cs.ucr.edu

Cormac J. Sreenan
Dept. of Computer Science,
University College Cork,
Ireland
cjs@cs.ucc.ie

This publication has emanated from research conducted with the financial support of Science Foundation Ireland (SFI) under Grant Number 13/IA/1892



Introduction

- Adaptive Bitrate Streaming (ABR) over HTTP is considered the default streaming approach for many video providers, such as Netflix, Hulu, etc.
- Driven by the availability of network infrastructure and the ability of HTTP to bypass firewalls
- Over 55% of mobile traffic is now video, and is expected to increase to 72% by 2019
- Video traffic accounts for a high portion of peak fixed and mobile access traffic
- High volumes in peak traffic forces a balance between maintaining stream quality and being fair to other users

Streaming Control Loop

- Video Providers (VP) manage the streaming control loop to improve client quality of experience (QoE)
- VP offer a number of representations, or quality levels, for each clip, or flow, thus providing a means of maximizing quality based on available bandwidth perceived by the client
- Furthermore each quality level is fragmented into numerous segments, or chunks, which allows the client to vary the achievable quality based on changes in the network
- These techniques offer a means of improving network utilization, while improving client QoE
- Clients maximize viewable quality in a greedy manner, irrespective of other users or underlying network conditions

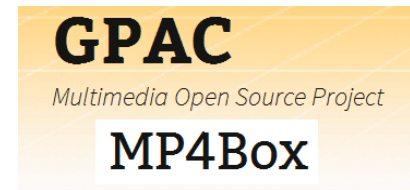
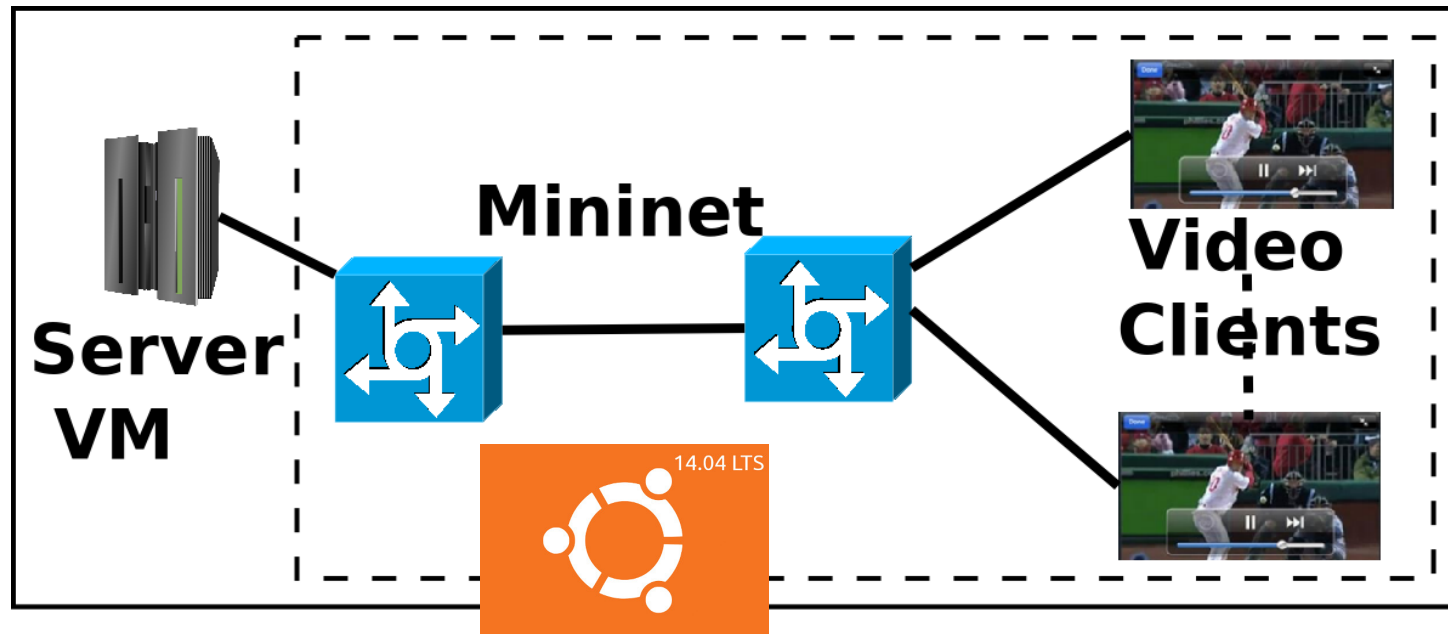
Network Control Loop

- Network Providers (NP) also employ several techniques to manage video traffic
- Caching provides a means of reducing network core bandwidth usage by offering popular content locally
- Transrating changes the encoding rate of a clip to suit current network conditions or reduce network load
- Traffic Shaping limits the video flow to a specific bandwidth level to ensure fair resource allocation
- These techniques improve the usage of network resources and ensure adequate services are available for all users

Motivation

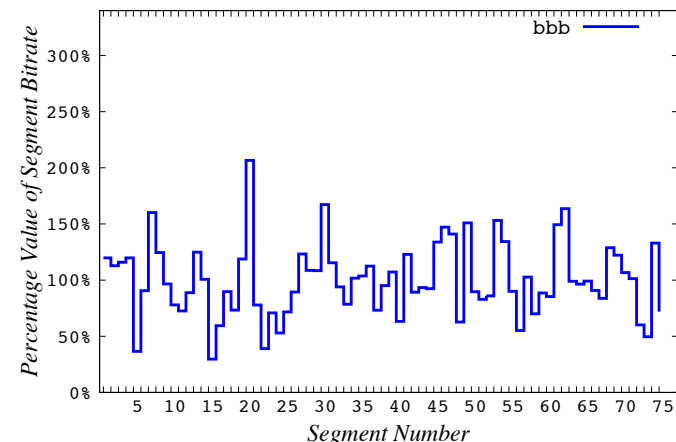
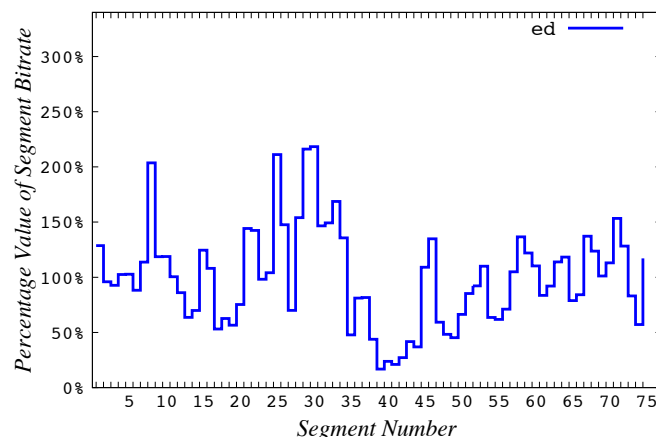
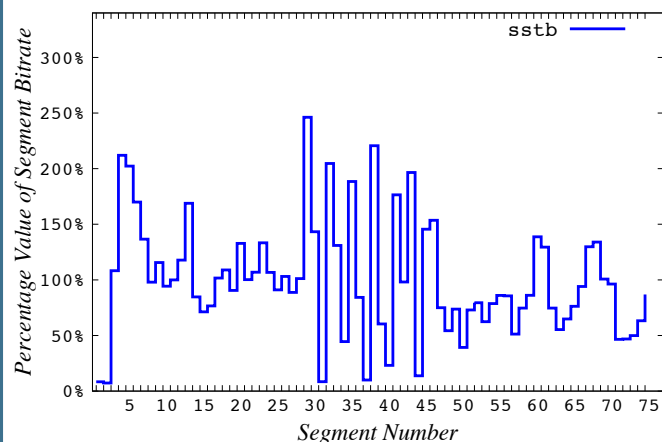
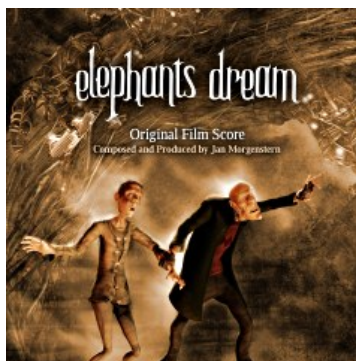
- Interaction between streaming and resource management control loops raise several design concerns
 - Will adaptation in one loop cause detrimental reaction in the other loop?
 - Can the network loop restrain the impact of the stream loop?
- In this work we focus on statically defining the bandwidth allocation and investigate the impact of traffic shaping on the performance of ABR streaming
- More specifically, we evaluate the impact of aggregate versus individual traffic shaping strategies over bottleneck links

Evaluation Test bed



- VM-based test bed utilizing an Apache server, Mininet and Ubuntu for the evaluation of DASH H264 encoded video streaming using the GPAC player MP4Box
- Nine distinct representation (quality) levels: 6Mbps, 5Mbps, 4Mbps, 3Mbps, 2.5Mbps, 2Mbps, 1.5Mbps, 1Mbps, and 0.5Mbps

Averaged Segment Rate Maps Per Clip



- Averaged segment rate maps, across all representation levels, for each of the evaluated clips: Sita Sings the Blues (SSTB), Elephant Dreams (ED) and Big Buck Bunny (BBB)
- The y-axis percentage value of segment bitrate offers a relative rate by which we can show the variability of the video content

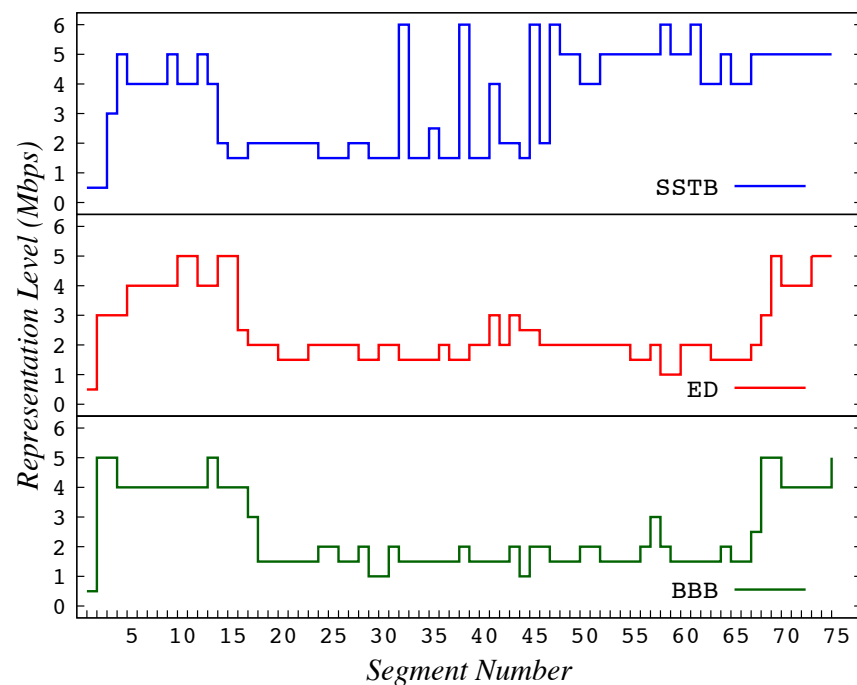
Evaluation Setup

- Aggregate case: six clients compete for a 30 Mbps link where the achievable bandwidth per segment of each client is dependent on the demands of the other clients
- Individual case: six clients share a 30 Mbps link where each client is assigned an equal share, 5Mbps, of the available bandwidth
- For each evaluation iteration, all three clips are viewed, with each clip being viewed by two clients, as two distinct streams
- The following evaluation results are typically reflective of multiple runs unless stipulated otherwise

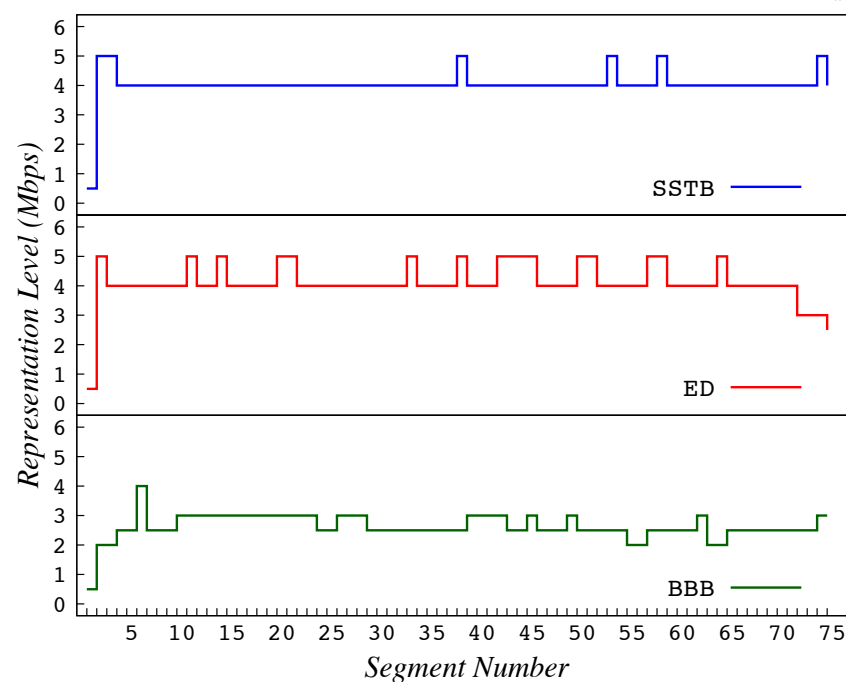
Variations in Achievable Quality



MISL



Aggregate



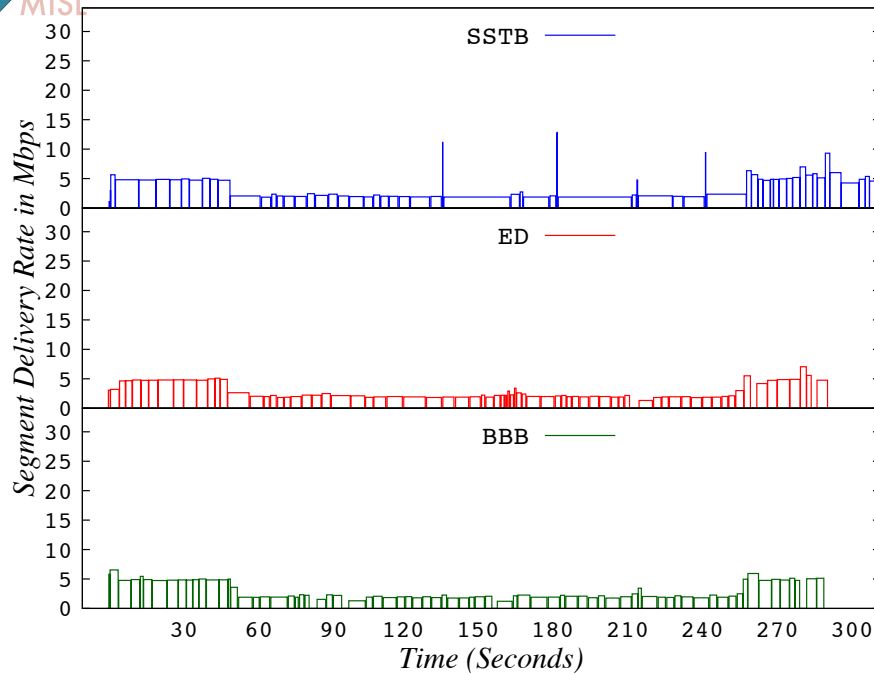
Individual

- Aggregate case: due to competition between clients, choosing the maximum achievable representation level in a greedy manner forces clients to make the wrong decision.
- Individual case: fixed bandwidth allocation removes competition and permits the client to gracefully move between quality levels, with a reduced level of variability ⁹

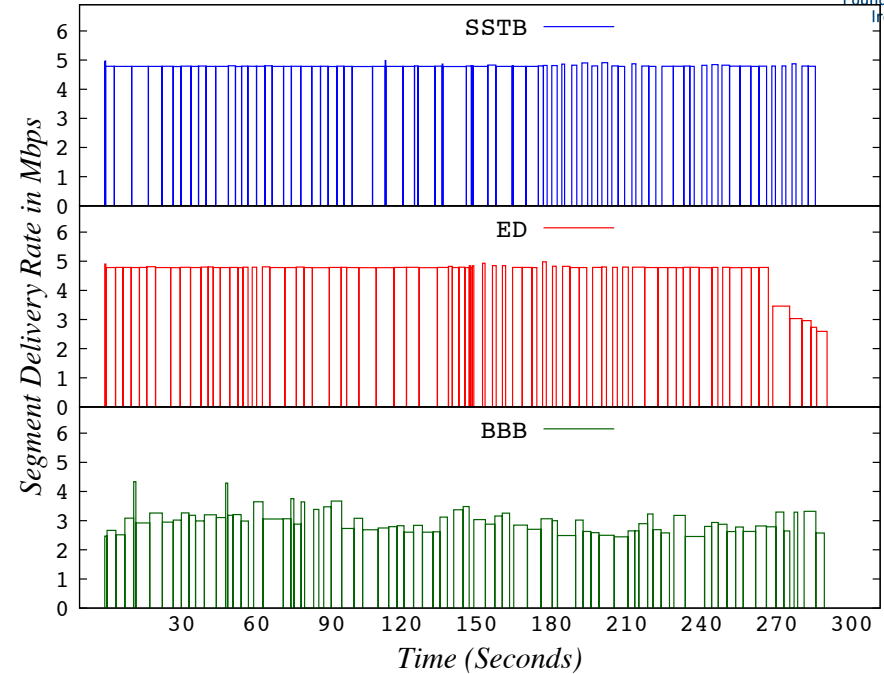
Segment Delivery Rate



MISI



Aggregate



Individual

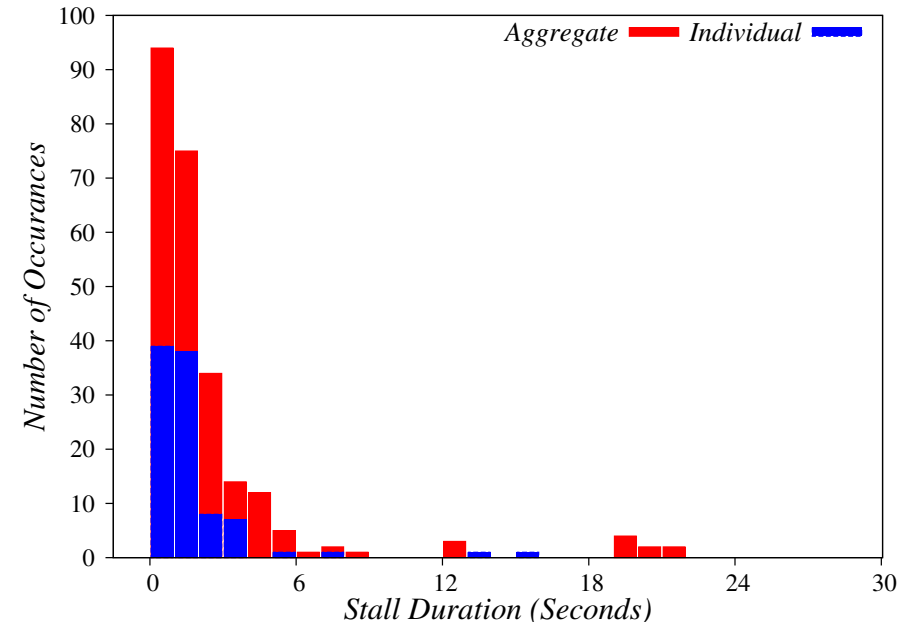
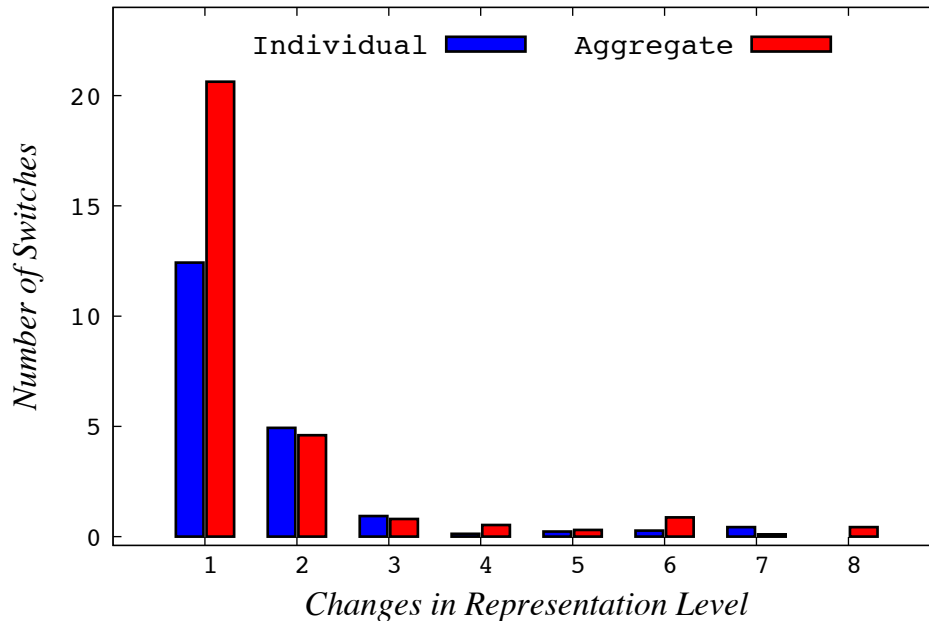
- Aggregate case: wrong choices impose delays in the delivery of segments, seen in the wide delivery time for some SSTB segments. As the segment duration is 4-seconds, dependent on buffer levels, any segment that take longer than 4-seconds to deliver could cause a stall.
- Individual case: narrow delivery times and little variation in delivery rate provides an improvement in consistency of quality



Switching Dynamics and Associated Stalls

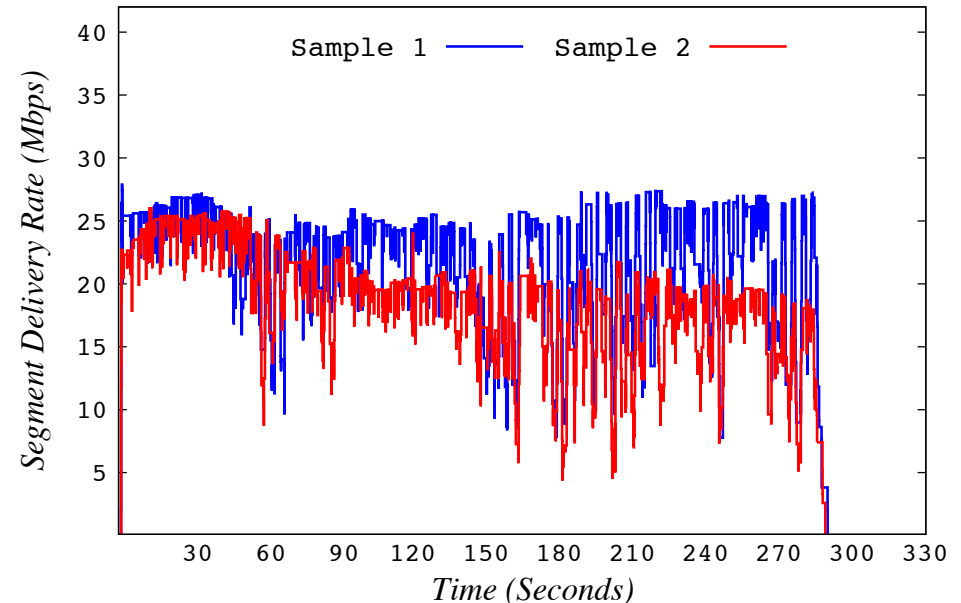
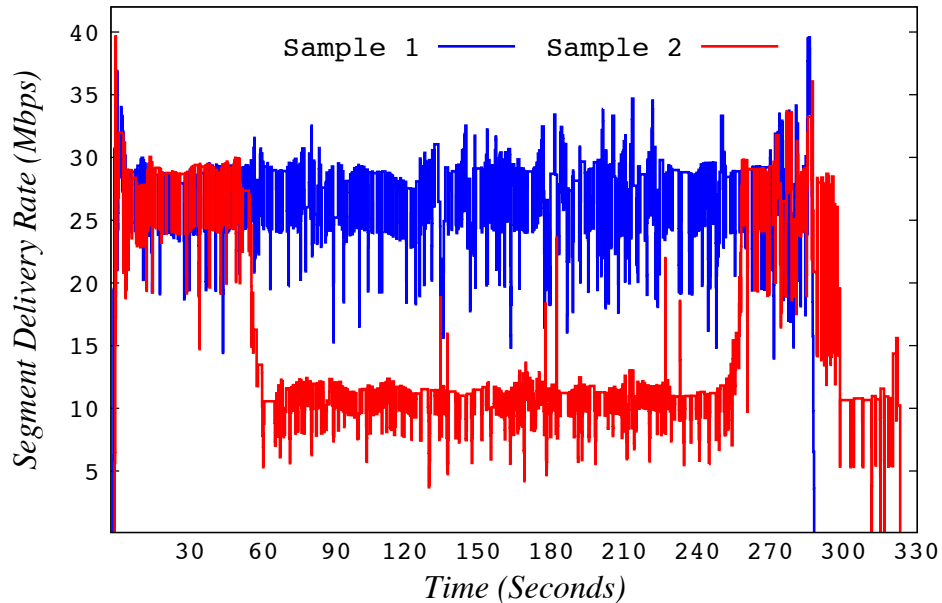


MISL



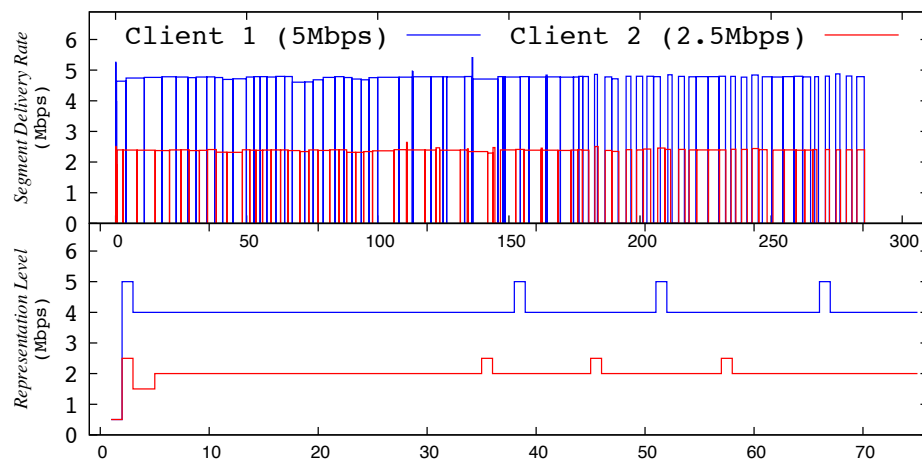
- Aggregate case: the wrong switching choice imposes a greater number of stalls and associated stall durations
- Individual case: imposing a fix limit on the bandwidth allocation reduces the number of switches and stall occurrences
- The wrong choices can be seen in both the switch dynamics (too many changes in quality) and stall durations and number of stall occurrences

Bandwidth Utilization

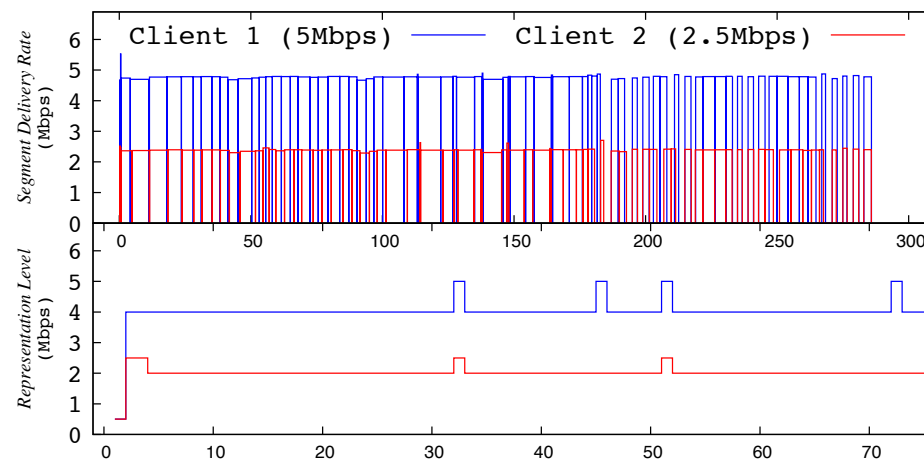


- Aggregate case: clients compete aggressively for maximum achievable quality, resulting in the wrong decision for client and network
- Individual case: clients provide greater levels of bandwidth utilization
- Similar average fairness values were seen for the aggregate and individual cases, but significant variations is seen in the aggregate case, while the individual case is more consistent across multiple runs.

Allocating Bandwidth based on Link Conditions



Standard MPD file



Modified MPD file

- Segment Delivery Rate versus Achievable Representation Quality level
- Evaluated based on clients with different Link Capacities (5Mbps and 2.5 Mbps)
- Streaming based on using the standard MPD file versus using a Modified MPD file which removes all representation levels above the Link Capacity allocated to the client

Conclusion

- In our work we show that traffic shaping can balance design objectives such as quality, fairness and network utilization
- As can be seen, we have highlighted a number of issues when the wrong choice is taken by the client
- Our work is about assisting the client to make the right choice for both streaming quality and resource management
- Using Mininet and ABR we showed that traffic shaping can smooth delivery and reduce peaks in network demand
- And does not cause an adverse interaction in the control loop between the ABR client and the video provider

Open Questions

- Determine results based on different streaming adaptation strategies, such as buffer-based estimation
- The design of the traffic shaping strategy is an interesting problem in itself
- When should traffic shaping start, such as from the beginning of the stream or after play out begins
- Should the bandwidth allocation rate not only consider resource availability but also encoding rate
- Can the network effectively control the streaming loop if the behavior of the client is known

Delivery of Adaptive Bit Rate Video: Balancing Fairness, Efficiency and Quality



Thank you for you time.

Questions?

