ABR Video Shaping

Why Mobile Operators Tend to Throttle Video

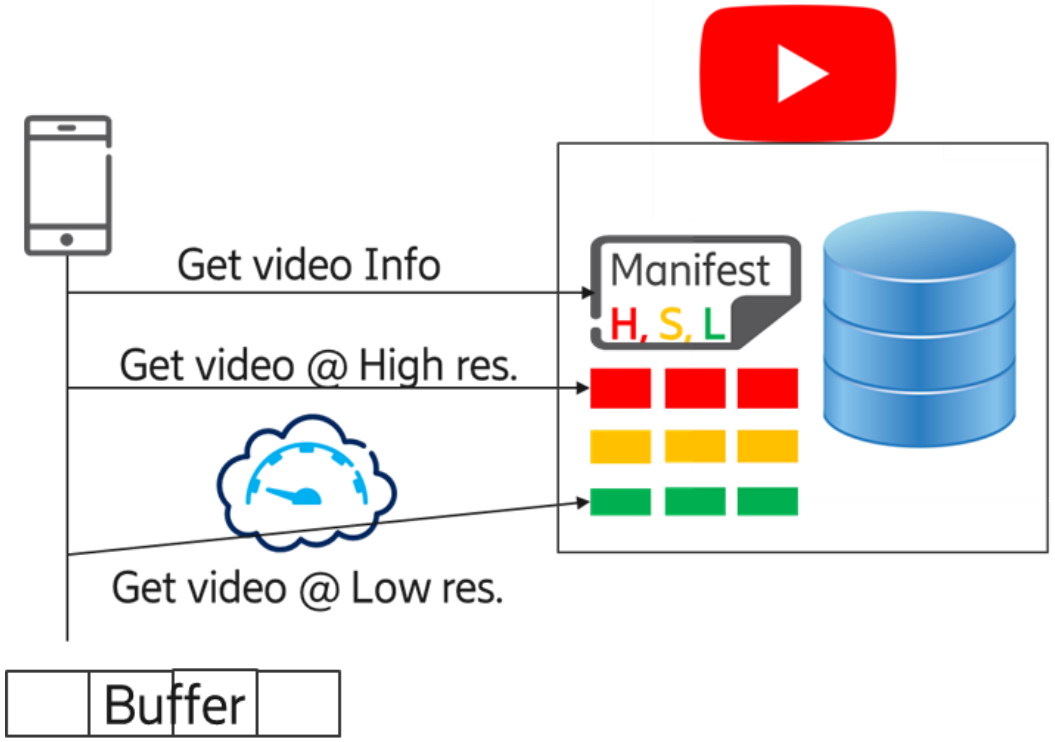
IETF 117
SADCDN Side Meeting

Overall objective of video shaping

- Problem:
  - The monthly data caps of mobile subscriptions are increasing, and it is getting more common with "unlimited data" subscriptions. Less incentive to "conserve data".
  - Use of video in social media applications is increasing (e.g., Tiktok, Instagram Reels, Twitter videos).
  - The rate of volume increase observed in networks is often faster than the rate of capacity upgrades.
  - Cell congestion is often transient and irregular, limited resources are shared among multiple users. Users in poor RF conditions can cause sudden capacity drops due to Proportional Fair scheduling algorithms.

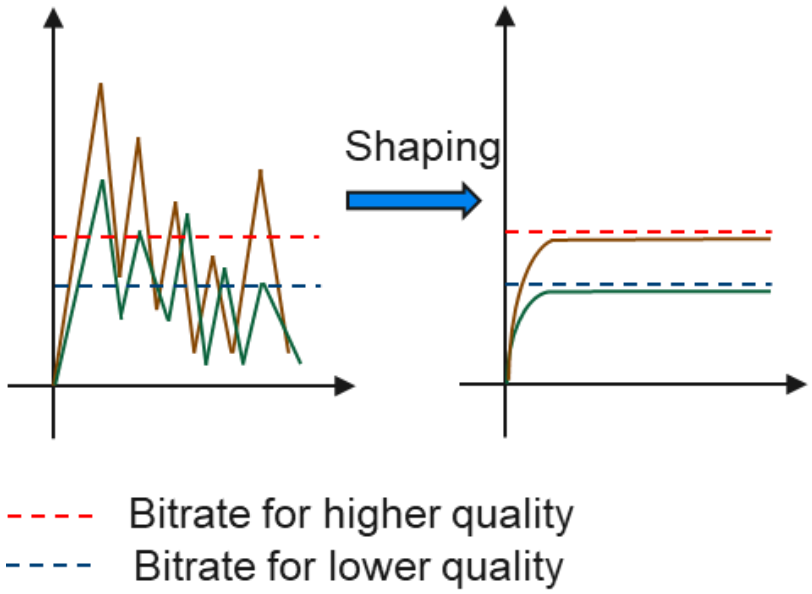
- Video throttling is applied with the intention to:
  - More efficiently use limited shared network resources.
  - Reduce traffic volume required for video to distribute capacity more evenly.
  - Provide a smoother playback experience for users when network conditions are highly variable.

Adaptive Bitrate (ABR) Video – Simplified



- Dynamic video resolution based on estimates and prediction of network capacity.
- Video divided into segments; each segment available in multiple resolutions and video qualities.
- Client selects resolution of next segment based on estimated network capacity and size of its playout buffer.

ABR Video Shaping



- Mobile network operators make use of ABR to reduce volume in their networks.
- Throttling is applied selectively to flows that are classified as ABR video.
- A bitrate is enforced such that the client is likely to select a video resolution that is lower than some quality threshold deemed “good enough” for common mobile clients.
- The desired effect is reduced cell load and improved Data Radio Bearer (DRB) throughput in cells that are moderately to highly loaded.

Issues with Video Shaping

- There are multiple vendors of video shaping solutions, each with slightly different ways of capping video resolution.
  - Shapers – delay and buffer packets when bitrate exceeds threshold.
  - Policers – drop packets when bitrate exceeds threshold.
  - APIs for self regulation – applications interact with networks through APIs where policies are expressed.
- Shaper and policer implementations vary in how they allow traffic bursts.
  - One or two colour token buckets
  - "Radio friendly pacing"
  - Etc.
- Operators need to configure bitrates that have the desired resolution capping effect while having minimal side-effects.
  - Video segments are not of uniform size, the act of shaping could lead to unwanted quality fluctuations.
- Current APIs for self regulation are complex and costly to implement.

What about 5G QoS?

- 5G Stand Alone introduces UE Route Selection Policies (URSP) where traffic receives different QoS depending on categories defined by the Operating System.
- Applications select most appropriate category, depending on its needs. Each category has a set of trade-offs:
  - High bitrate, prioritized below other categories.
  - Lower average bitrates, prioritized over high bitrate.
  - Etc..
- With better QoS control it is possible to steer bandwidth hungry traffic to lower priority bearers and prioritize other types of traffic when resources are scarce. This can potentially reduce the need for video shaping.