HIGH EFFICIENCY STREAMING

Fast Zapping

Whitepaper

HESP ALLIANCE

High Efficiency Streaming

Fast Zapping

Operators are increasingly looking into transitioning from a parallel broadcast digital TV / IP television (IPTV) and over-the-top (OTT) workflow to a next-generation unified OTT workflow.

However, they cannot deliver upon the increasing OTT viewer Quality of Experience (QoE) expectations of low latency, to avoid spoiler effects from social media and neighbors, and fast zapping, which are crucial for a leanback TV experience.

High efficiency streaming will transform the industry as it brings OTT channel switching times on par with analogue television while immediately starting at the correct live latency. Whereas for IPTV substantial investments have to be made to achieve fast zapping, high efficiency streaming will inherently enable this for OTT, bringing important cost savings to operators.

Moreover, fast zapping will also create true multiview experiences as OTT viewers will be able to very quickly switch between different camera angles of a live event.

Introduction to High Efficiency Streaming

Today streaming protocols are typically classified through the latency they bring: the time it takes between the moment that action happens and the moment that a viewer sees this action on his screen. This latency can vary from more than 60 seconds in OTT use cases to sub second for video conferencing and other similar use cases. Specifically for live and interactive events it is key to have this latency as low as possible. Another criteria often used to classify streaming protocols is the ease of scalability. Protocols such as WebRTC can provide for sub-second latency, but are complex and expensive to scale as each client requires a persistent connection with the backend. HTTP based streaming protocols, on the other hand, are very scalable, through commoditized CDNs.

Viewers are more demanding than ever and want high efficiency streaming solutions. They do not just expect sub-second latency at scale. They expect a leanback TV experience over OTT. Viewers want the flexibility and universal device reach offered by OTT video but with low latency, to avoid spoiler effects from social media and neighbors, high video quality without loading times and fast zapping, on par or faster than the zapping experience they have grown to love from the broadcast environment.

Also operators are increasingly looking into high efficiency streaming as they want to transition from a parallel broadcast digital TV / IPTV and OTT workflow into one unified OTT workflow, to deliver upon the increasing viewer expectations, and to achieve cost savings by having only one workflow to buy and maintain. Such a transition to high efficiency streaming will not only bring cost savings to operators but will also make it possible to create new revenue streams through, for example, more targeted ad insertion. However, today operators cannot provide for high efficiency streaming with a unified OTT delivery solution, as leanback TV experiences cannot be realised through protocols such as WebRTC or legacy DASH and HLS, and their low latency implementations.

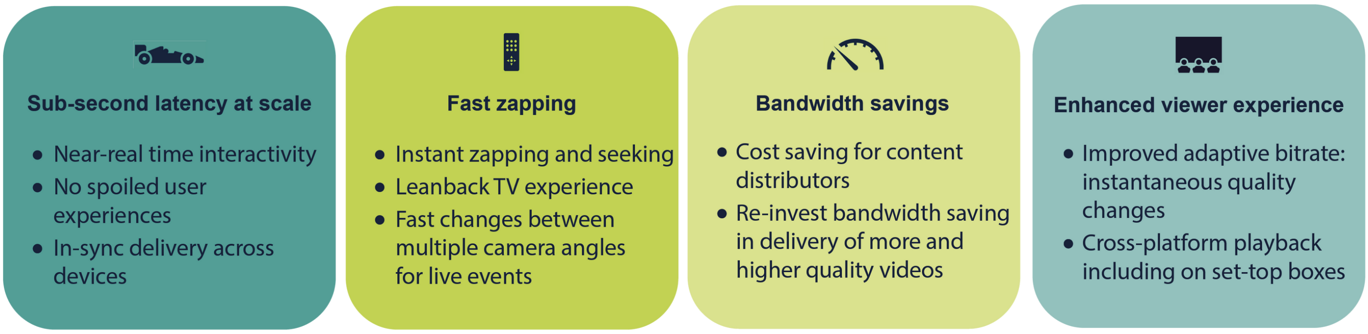

Figure 1: Increasing viewer QoE expectations require high efficiency streaming solutions.

This is where the High Efficiency Streaming Protocol (HESP), an HTTP based streaming protocol for next generation services, comes in as it provides for high efficiency streaming to deliver upon the increasing viewer expectations. In addition to sub-second latency at scale, it brings benefits such as:

- Improved adaptive bitrate (ABR);
- Bandwidth savings, up to 20%, to further drive cost savings for operators, and;
- Fast zapping, as low as 100ms, which makes it on par with analogue zapping.

In this whitepaper we will focus on fast zapping through high efficiency streaming with HESP.

Why fast zapping is important

Channel switching times are a critical QoE metric for video delivery systems. A study from Akamai has shown that people tend to abandon video services rapidly after only two seconds of delay. As a comparison, legacy analogue TV has channel switching times in the 0.1s to 1s range. When looking at IPTV in comparison with analogue distribution, channel switching times have increased significantly which resulted in substantial investments in IPTV by operators to reduce channel switching times. More specifically, there are server-side software solutions on the market to reduce channel switching times for IPTV, for example the Synamedia Visual Quality of Experience (VQE) technology. Similarly, operators have also invested in digital cable TV to reduce channel switching times through the use of more expensive set top boxes with additional tuners, to anticipate channel changes.

Viewers also expect this fast zapping experience for OTT video. However, OTT video delivery suffers from significantly higher channel switching times which can go up to multiple seconds. With many operators merging their existing parallel IPTV and OTT siloed solutions into a unified next-generation OTT delivery solution, it is important to maintain the quality of experience that online viewers are used to, including instant start-up, channel switching and seeking times, both on the big screen and on other consumer devices. Video content brings in viewers, but the quality of experience retains them.

Fast zapping can be a powerful tool for all types of interactive experiences, especially when paired with the sub second latency that high efficiency streaming brings. Imagine an event being recorded from multiple camera angles.

As users would be able to swiftly and efficiently switch between angles, a true multiview experience could be established where a viewer can select any camera angle and is able to instantly get in on the action from this new perspective. Switching viewing angles could be extremely responsive, with immediate feedback at the press of a button or a swipe on the screen.

The below picture shows such an HESP proof of concept which allows viewers to easily swipe using their phone or tablet from one camera angle to the other. Camera angle switching times measured vary from 150-400ms.

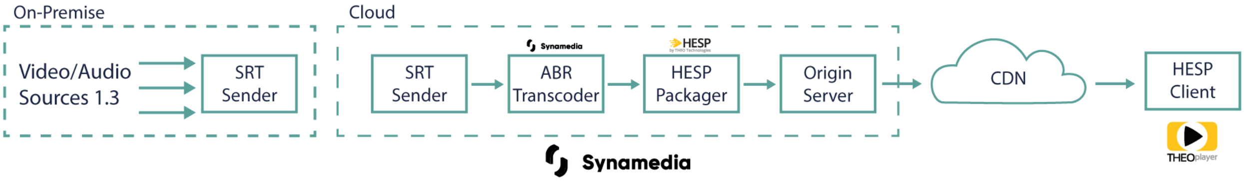

Figure 2: Setup to demonstrate fast camera angle switching times of 150-400ms .

Why channel switching times for OTT video are a challenge

With the introduction of OTT video, switching between different channels becomes noticeable. Channel switching times have improved over time, but popular streaming protocols like HLS and DASH still introduce intrusive and noticeable delays between switching channels. The reason for this is that distribution approaches such as HLS and DASH require a trade-off between live latency and channel switching time. When live latency is not a focus item for legacy HLS and DASH, it is possible to fully focus on channel switching time.

At any moment in time the client can request a segment and start playback from the first IDR frame of that segment. When starting playback as soon as there are a limited number of frames in the buffer, channel switching time can be very low.

However, when live latency is critical, there is a trade-off to be made. This is also the case for the low latency implementations of HLS and DASH where significant channel switching times can be seen, which can go up to two to three seconds. As an example, using LL-DASH you can choose for fast start-up. This means that you start playback from the last segment in the manifest.

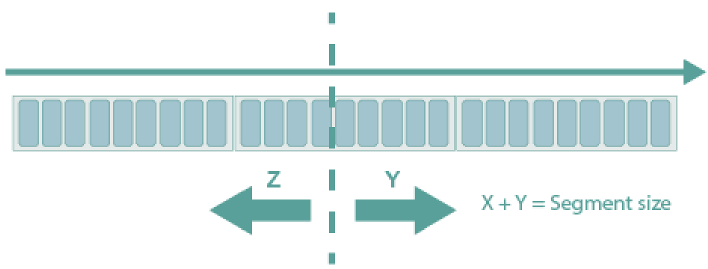

Figure 3: LL-DASH brings a trade off: either there is X live latency or Y channel zapping latency, or a combination.

As LL-DASH requires that you always start playback from the start of the segment, it implies that in most of the cases you do not start playback at the correct live latency: if the live edge is already 5 seconds within the 6-second segment then the live latency will start at 5 seconds instead of the target live latency that may be only 2 seconds. In order to get to the target live latency the client will be forced to play the content at a speed that is slightly higher than real-time to catch up on live latency. If on the other hand you want to start playback immediately at the right live latency then in some cases the client will need to wait until the next segment becomes available (providing the required IDR frame) which will have a negative impact on the channel switching time.

Similarly, in order to get a low channel switching time with LL-HLS, playback must start from a partial segment (part). The challenge with LL-HLS is that not every part must start with an IDR frame (or contain an IDR frame). This means that the client may not be able to start playback from the live edge. In this case the player again has to make the trade-off between channel switching time, and start up with an incorrect live latency, or start playback with the correct target live latency, in which case the channel switching time will be suboptimal.

Existing work arounds for OTT video have their limits

Current streaming protocols, such as LL-HLS and LL-DASH, but also WebRTC, have to make a trade-off between live latency and channel switching time. When starting at a higher latency, it is possible to catch up on live latency by playing back faster than real-time, however, this is well noticeable for certain audio content, and definitely for music that you know well.

As an alternative, the first segment(s) could be downloaded at lower quality/bitrate to speed up the download, but of course this comes at the expense of quality.

A last solution would be to download the entire segment and decode faster than real time in order to start at a lower latency, but when bandwidth isn't ample, this will increase channel switching time significantly.

Fast zapping with HESP

Instead of using a segment-based approach, HESP leverages a frame-based streaming approach, which does not require a trade-off between live latency and channel switching time. More specifically, HESP uses two streams:

1. An initialization stream, which contains only key frames. This stream is not regularly used. It is only used when we start a new stream.

2. A continuation stream, which is a regularly encoded stream for low latency purposes, which can continue playback after any initialization stream image.

This two-stream approach ensures that it is possible to always start playback at the most recent position in the video stream, and that a viewer can instantaneously start streams or zap to another stream at any given moment.

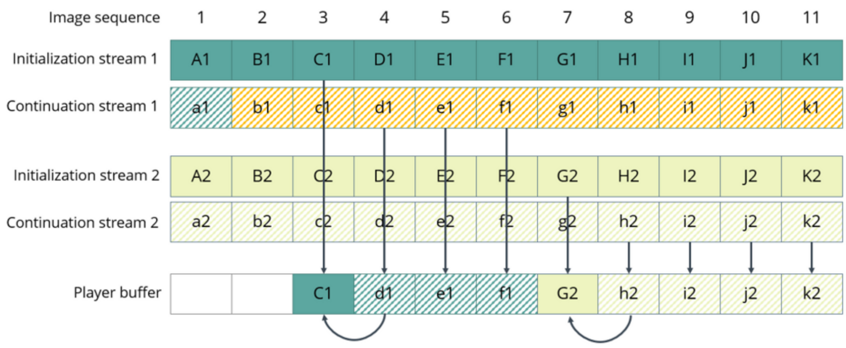

Figure 4: HESP uses two complementary streams. Whenever a user wants to start a new video, firstly an image or frame is fetched form the initialization stream. Images can be requested at any moment to start playback. Subsequently, images are fetched from the continuation stream. The continuation stream can playback at live latency after any initialization stream image.

When a channel switch happens, the player will first load an image from the initialization stream. As the initialization stream's images are all key frames, playback can start immediately. Key frames are expensive in terms of bandwidth, so we do not want to continue playing out images from the initialization stream. This is where the continuation stream kicks in. After the key frame of the initialization stream, images are requested from the continuation stream, starting at exactly the right location by using a byte range request.

The result when combining these two actions is straightforward: the player will instantly get all data to start playback at the current live point without any trade-off in latency. And also without investments in separate fast zapping solutions. HESP inherently has fast zapping capabilities and hence brings substantial cost savings.

Interested in joining the HESP Alliance?

Contact us.

HESP ALLIANCE