Myelin Whitepaper Series

LATENCY IN ONLINE STREAMING

Impact & Solutions

7th August 2023

Mayank Shukla

Product Marketing, Media & Entertainment

Myelin Foundry®

Everything AI, Anything Creative

TABLE OF CONTENTS

INTRODUCTION

As video streaming continues to gain popularity, there is a growing need to overcome challenges and enhance the user viewing experience for optimal quality. These challenges must be addressed with more people watching movies, TV shows, and other videos online.

The most pertinent of those is perhaps the issue of latency which has been a hurdle in providing a seamless viewing experience for users. Several factors contribute to latency, such as the distance between the server and the user, the quality of bandwidth for internet connectivity, and the amount of traffic generated over the network. This comes with the need to deliver high-quality video content, as anything less would be jarring for users, ultimately driving them away from the streaming platform.

We will discuss this latency problem in depth to analyse its effects as we advance. Also, there would be a discussion about the various solutions through which attempts have been made to mitigate this issue, primarily through AI-based solutions. It would involve the extent to which the technology can be used to optimize the delivery of video content to users based on their location and internet connection speed. AI can also be used to improve video compression quality, which can help reduce latency and bandwidth requirements.

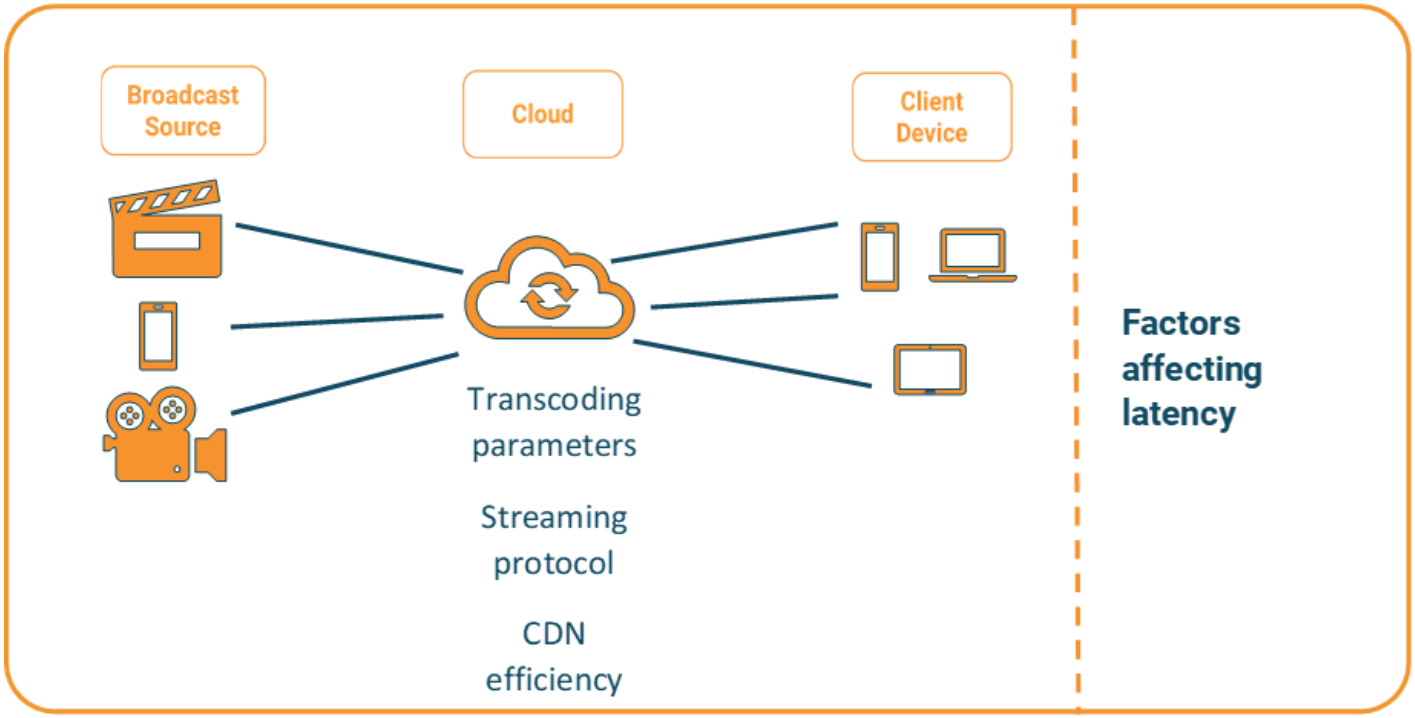
The future of video streaming looks bright. As AI continues to develop, we can expect to see even more improvements in the quality and reliability of video streaming. This will make it possible for people to watch their favourite videos anywhere, at any time, without worrying about buffering or choppy playback.

LATENCY

Latency in video streaming refers to the delay between when an event occurs, captured live on video camera, and when it is displayed on the user's device during live streaming. It is calculated as the time it takes for a data packet to travel from the video source to the viewer's device. This is commonly measured in milliseconds (ms).

The amount of latency can vary depending on several factors, including the encoding and decoding process, network speed, and the distance between the user and the streaming server. Latency can impact the synchronization of audio and video, leading to a poor viewing experience for users. Various types of latency can be defined as:

- **Transmission Latency:** Time taken for a data packet to travel from the video source to the viewer's device. Measured in seconds, it can vary depending on the distance between the source and destination and the number of network hops along the path.
- **Processing Latency:** The processing latency on the viewer's device can include encoding, decoding, compressing, decompressing, and buffering data. It can vary depending on the device's processing speed and the software's complexity.
- **Streaming latency:** The time it takes for the video stream to be delivered to the client device which is determined by the network bandwidth and the streaming protocol.
- **Playback Latency:** It is the time taken for the video to be decoded and displayed on the viewer's device. Playback latency can be affected by a variety of factors, including network congestion, buffering, and the processing power of the viewer's device.



UNDERSTANDING WHAT CAUSES LATENCY

Streaming video content involves various technical processes that can affect the user's experience, particularly regarding latency or delay between the video source and the viewer. These processes include encoding and packaging the video content, uploading it to a content delivery network (CDN), propagating the content across different caches, delivering it to the user's device, and buffering the video in the player.

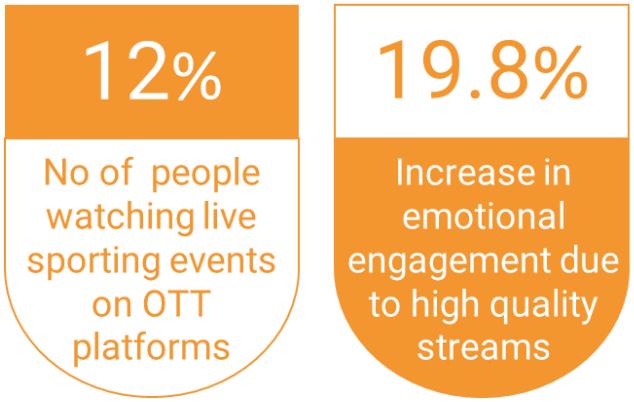
The total latency of a video streaming session is the sum of all these individual latencies. To achieve low-latency video streaming, minimizing each of these latencies is important. Uploading the packaged content to a CDN can also impact latency, especially over a wireless connection.

Content delivery networks frequently deliver content at scale, which can introduce additional latency as content propagates between different aches. The user's network connection, such as connecting to a Wi-Fi hotspot or using a mobile connection, also impacts latency. The player buffer is another component that can affect latency, as video players need to buffer media to ensure smooth playback.

Optimizing these technical processes can go a long way in improving the streaming experience and minimizing latency issues.

LATENCY IN DIFFERENT KINDS OF CONTENT

Before we proceed, let us see a basic explanation of real-time streaming. It refers to delivering synchronized content to many viewers, with a delivery time of less than 1 second per user.



Live video streaming has become an increasingly popular method of consuming media content in recent years. With technological advancements and internet infrastructure, users can now access high-quality live video streams from anywhere globally. However, one of the primary challenges of live video streaming is the issue of latency and rebuffering, which can significantly impact the user experience. Here is how different types of content can be affected by latency and resulting in losses for platforms:

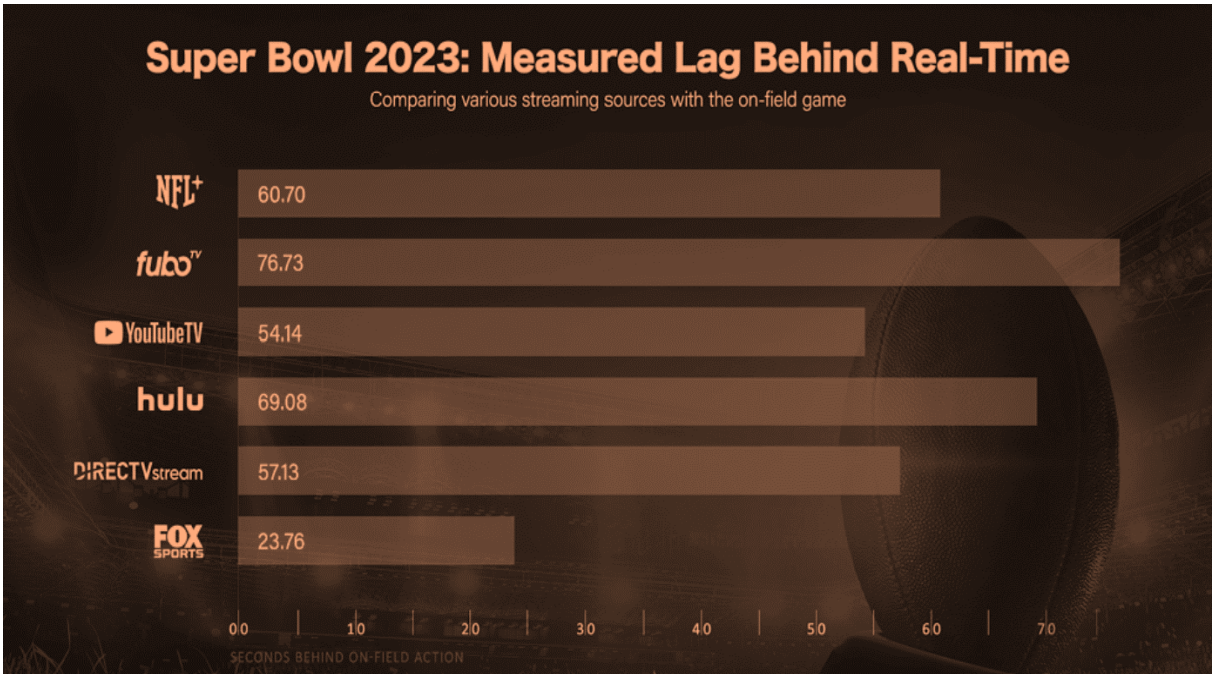
SPORTS CONTENT

Live streaming is one of the most latency-sensitive types of content, as viewers expect to see events in real-time. Even a delay of a few seconds can significantly impact the user experience, as it can cause viewers to miss key moments or lose the sense of immersion and engagement that is central to the live experience. High latency in live streaming can result in reduced viewership, lower engagement rates, and increased user frustration, all of which can lead to lost revenue and brand damage for platforms.

28% Estimated Pay-Tv market share of global sports fans by 2026

One use case that comes to mind when low-latency streaming is discussed is sports. Super Bowl LVII took place in Glendale, Ariz., with an estimated audience of 130 million worldwide.

Some estimates say that 7 million viewed the big game via streaming, customers may sometimes watch on multiple devices.



As per a report by Phenix, there were delays going beyond 60 seconds during the live streaming of SuperBowl 2023.The user reactions responding to such delays were also recorded and they were like:

"Talking to my father while he watched the same over cable-TV network- about 20 seconds delay for my IPTV service"

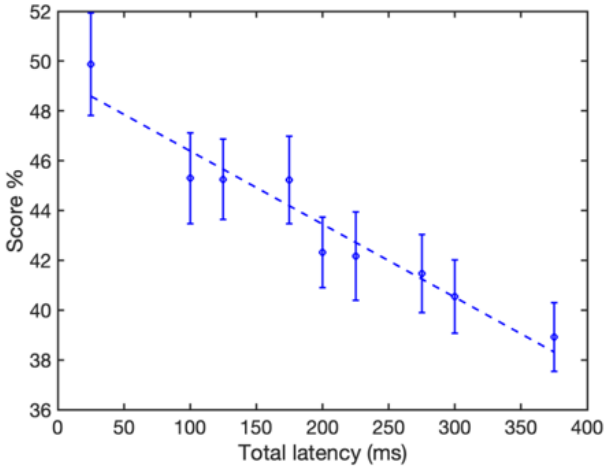
"Saw a live tweet for a massive play when I hadn't seen the setup yet"

"Friend messaged me that my team just scored, about one minute before viewing it"

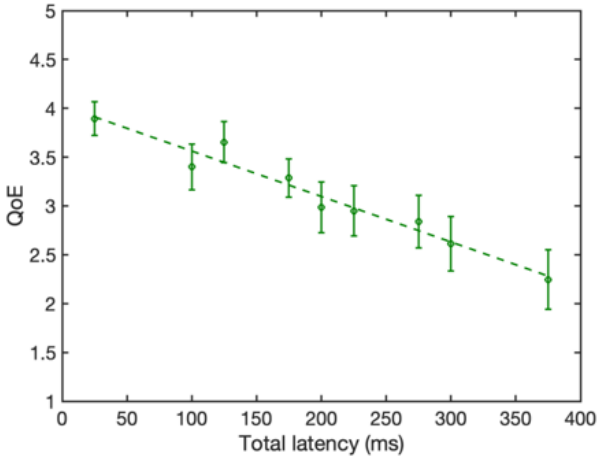
GAMING

Gaming is perhaps the most latency-sensitive type of content, as even small delays can significantly impact gameplay and user experience. High latency in gaming can result in slow response times, lag, and glitches, all of which can lead to lost games, reduced user satisfaction, and increased user churn. For gaming platforms, high latency can result in reduced user engagement, lower player retention rates, and negative word-of-mouth, all of which can impact revenue and brand reputation.

According to a study conducted in 2022 to find out the impact of latency in an online game, it was seen that latency had some impact on the player's overall gaming experience. Its impact can be observed in terms of the player performance and overall Quality of Experience (QoE) as indicated by the graphs below:



Player Score v/s Latency



QoE v/s Latency

USER-GENERATED CONTENT

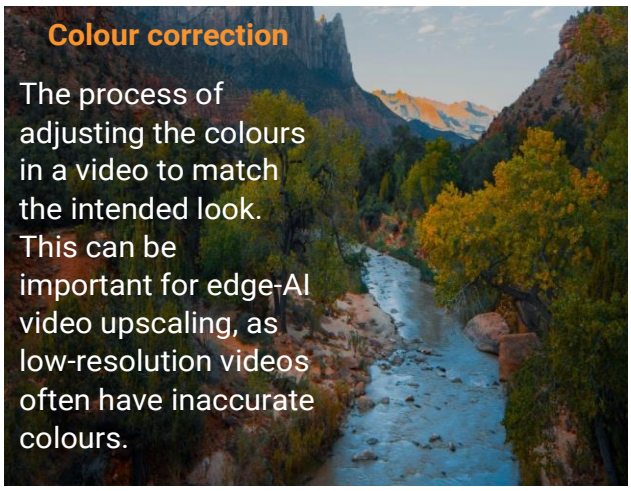
Live user-generated content (UGC) is another type of content that can be impacted by latency, particularly if it involves real-time interaction or collaboration between users. For example, video conferencing and social media platforms rely on low latency to support real-time communication and engagement between users. High latency in UGC can result in poor audio and video quality, disrupted communication, and frustrated users, all of which can lead to lower user satisfaction and reduced usage of the platform.

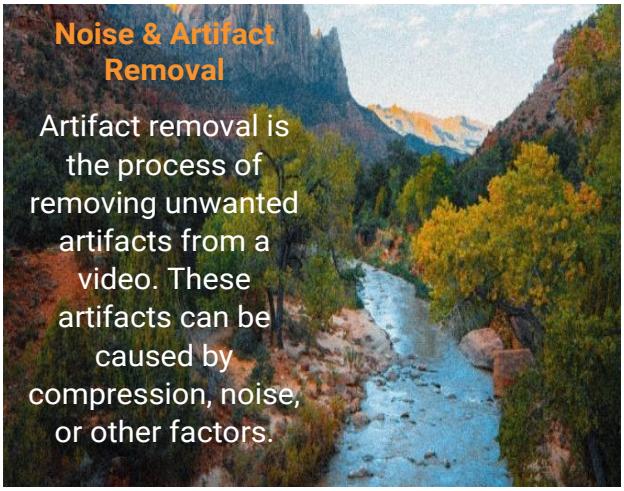
With the kind of content generated across the social media platforms and its monetization through various, there is need to ensure that best content quality is provided in real time to gain more viewer base for the good content creators.

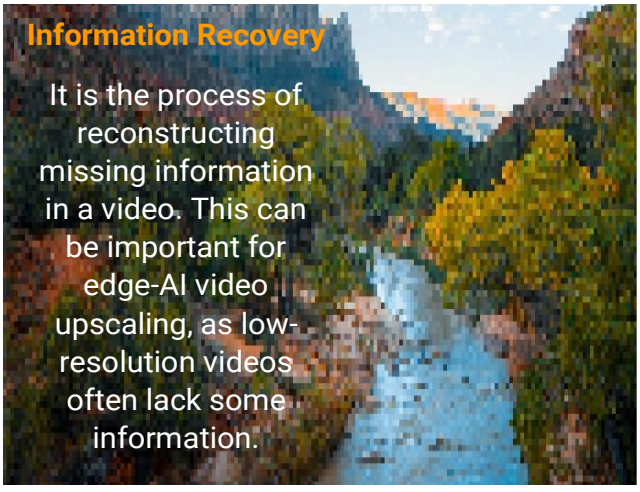
VIDEO UPSCALING AND EDGE-AI

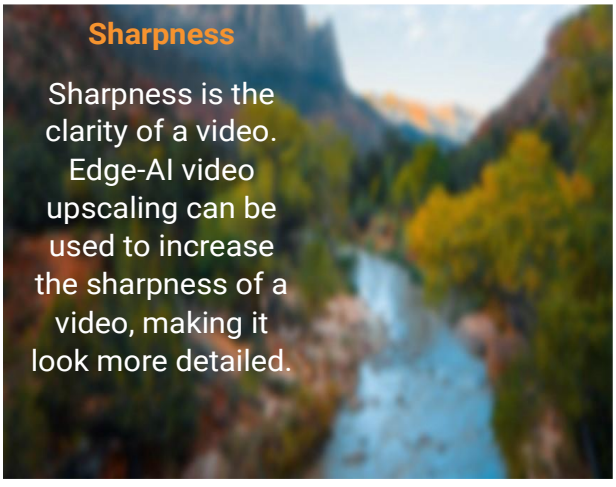
Video upscaling is the process of increasing the resolution of a video and enhancing the perceptual quality. Video upscaling through the Deep learning models is a powerful method for perceptual video quality enhancement. These models are trained on a large dataset of high-resolution and low-resolution videos, where they learn to identify and reconstruct the missing information in low-resolution videos while reducing unpleasant perceptual artifacts. Once the deep learning models are trained, they can be used to upscale videos in real time. In addition to improving video quality, edge-AI video upscaling can also be used to reduce latency during live streaming. This is done by performing the upscaling process on the edge device, rather than on a remote server. It reduces the amount of data that needs to be sent over the network, which ensures superior quality of viewing experience.

Edge-AI video upscaling can produce higher-quality results with fewer artifacts, can be more computationally efficient, which makes it ideal for applications where real-time performance is critical, such as live streaming. Also, it can be used to upscale videos on a variety of devices, including smartphones, tablets, and laptops.By incorporating the following techniques, edge-AI video upscaling can significantly improve the quality of low-resolution videos. The following components form a part of Myelin's FoveaStream edge-AI video upscaling its solution-

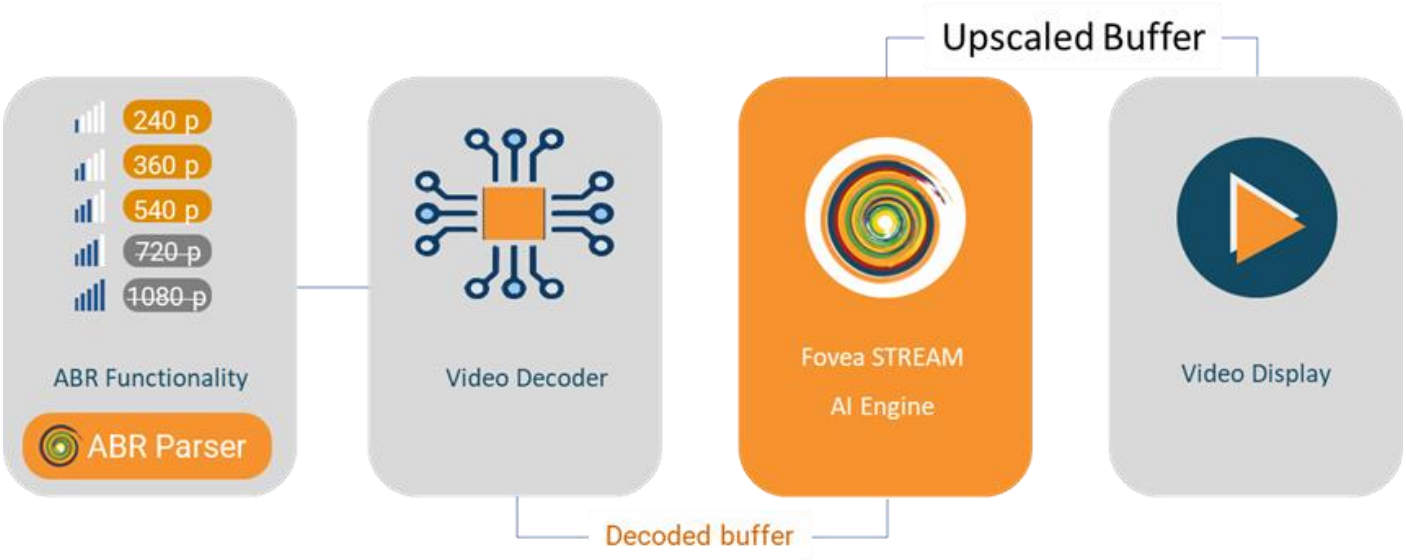




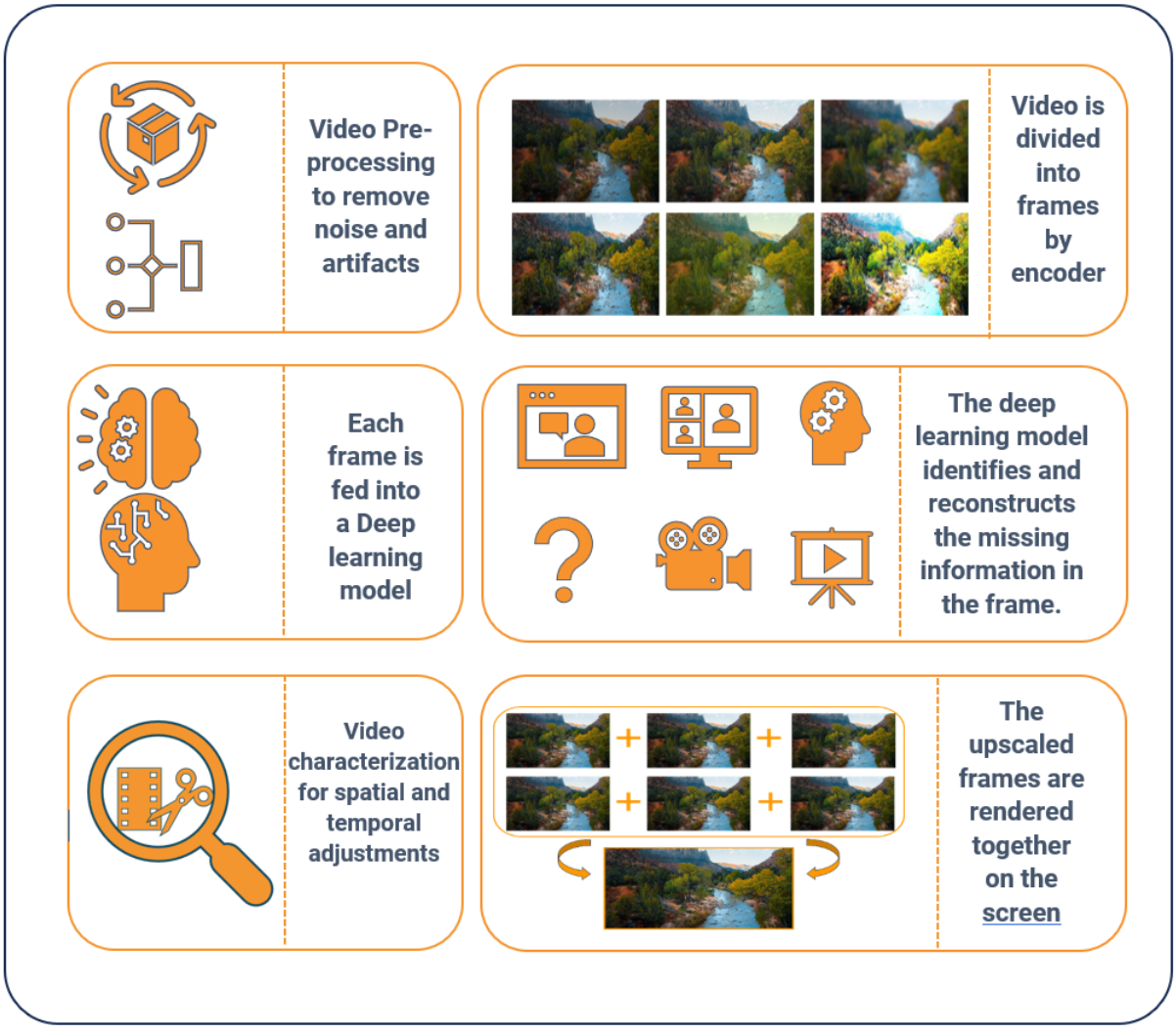


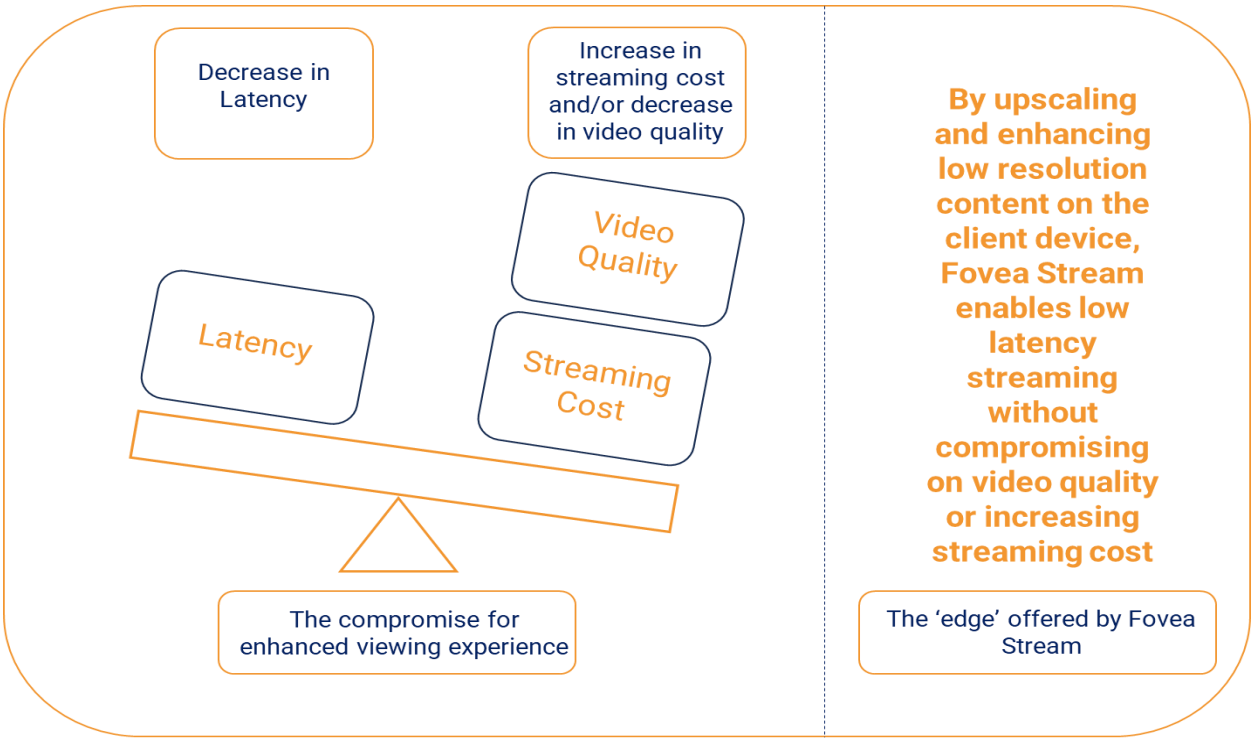


THE EDGE-AI MECHANISM



INSIDE THE AI-ENGINE





The table below shows comparison on various parameters while streaming SD content in parallel across Fovea Stream and YouTube simultaneously in the low latency mode.

Parameter \ Platform	FoveaStream	YouTube
Broadcast resolution	540p	540p
Broadcast bitrate	1.2Mbps	1.2Mbps
Streaming mode	NA	Low latency
Playback quality	FHD - 1080p	SD 480p
End to end latency(in sec)	6-8	9-11
Comparison with YouTube Low latency mode		

The total latency numbers shown in the table are a combination of broadcast device and playback latency. The foremost challenge while mitigating this latency issue is a quality compromise in the content for the viewers. Traditional solutions often trade quality for reduced latency as they go for reducing the bitrate of the video stream, which leads to a loss of detail and sharpness. The thing that sets apart Myelin's Fovea Stream is that it maintains a high quality FHD output even at low latency numbers. Also, due to its seamless client-side integration no modifications are required for existing streaming pipelines, thus ensuring no tampering with the existing workflows. It's critical because the entire framework is based on working on the edge device network.

CONCLUSION

Latency is a critical issue in video streaming due to its impact on viewing experience and can lead to reduced viewership, engagement, and revenue loss. Traditional solutions for low latency streaming often compromise video quality. By upscaling videos on edge devices through its Edge-AI technology, Fovea Stream maintains high-quality visuals, reduces latency, and lowers costs. It enhances video quality while intelligently optimizing the stream to retain sharpness and detail, without the need for extensive modifications to existing workflows.

Myelin's Fovea Stream offers a compelling value proposition for the rapidly growing video streaming industry. By combining high-quality streaming, low latency, and effortless integration, Fovea Stream stands out as a promising solution in the quest for an improved video streaming future. Fovea Stream's unique value lies in its ability to harmonize both aspects i.e., video quality improvement and latency reduction, positioning it as a game-changer in the evolving video streaming landscape. In addition to its high quality, Fovea Stream is also very scalable. It can be used to stream videos to many viewers without sacrificing quality or performance.

REFERENCES

1. THE FUTURE OF STREAMING IS FAN FIRST (2023, March) https://www.sportspromedia.com/wp-content/uploads/2023/03/Endeavor-Streaming-X-SportsPro-The-future-of-streaming-is-fan-first.pdf
2. The State of Real-Time Streaming 2022 (2022, March 26) Tim Siglin https://www.streamingmedia.com/Articles/Editorial/Featured-Articles/The-State-of-Real-Time-Streaming-2022-152089.aspx
3. The four elements of video performance (2016, August 18) Jon Dahl https://www.mux.com/blog/the-four-elements-of-video-performance
4. OTT Statistics For 2023 (2023, March 6) Daniel Ruby https://www.demandsage.com/ott-statistics/
5. 10 Streaming Service Statistics We're Watching in 2023 (2023, March 31) tvScientific https://www.tvscientific.com/insight/streaming-services-statistics
6. OTT Global Viewpoint (2018, October) Philip Hunter https://www.thebroadcastbridge.com/home/category/ott/entry/12088/ott-on-the-rise-but-still-plagued-with-buffering-and-high-latency
7. Measuring Video Quality and Performance: Best Practices (2020, May) Akamai https://www.akamai.com/site/it/documents/white-paper/measuring-video-quality-and-performance-best-practices.pdf
8. The Impact of Latency on Navigation in a First-Person Perspective Game (2022, April 29) Shengmei Liu, Mark Claypool: Worcester Polytechnic Institute - Association for Computing Machinery https://web.cs.wpi.edu/~claypool/papers/nav-chi-22/paper.pdf

ACKNOWLEDGMENT

- **Anusha Rammohan**, Senior Director Media & Entertainment
- **Mansi Vyas**, Regional Sales Head MENA & SEA
- **Dinesh Babu Sukumar**, Product Marketing Manager