# LCEVC x264 Report: Live Sports & eGames, ABR ladder

Jan Ozer, April 2020

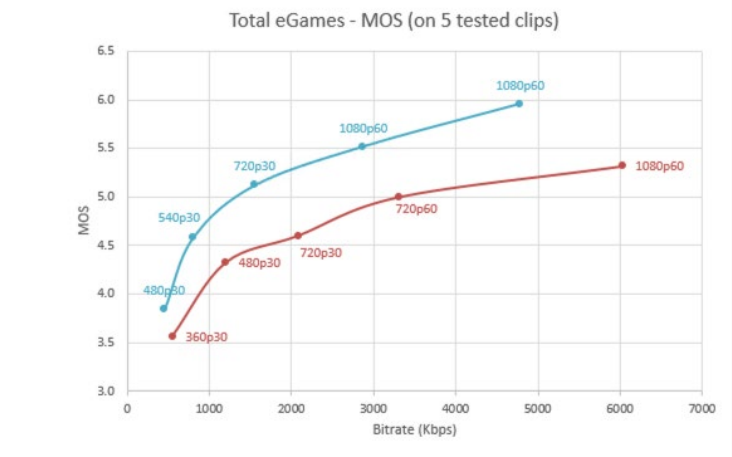
# Table of Contents

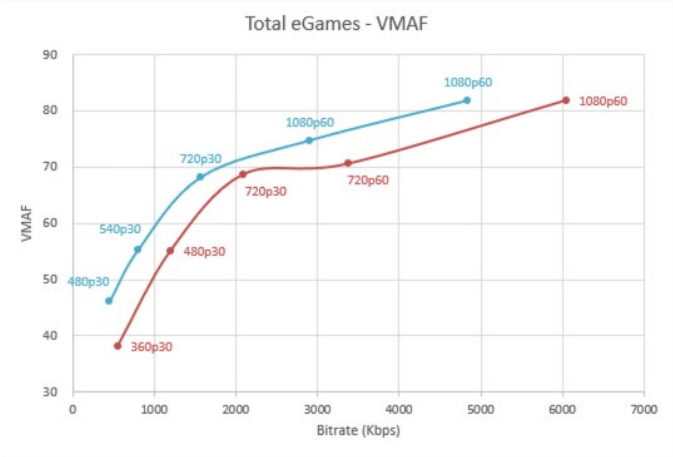
## Core Report (this document)



## Appendix (separate document)

# Overview




V-NOVA engaged Jan Ozer (dba Doceo Publishing) to evaluate the quality achievable with LCEVC enhancement when applied to multiple base-layer codecs.

The initial tests compared LCEVC with a base layer of x264 (LCEVC x264) to native x264. The first set of results, published in July 2020, **compared full resolution (1080p)** x264 encodes to LCEVC x264 encodes at lower data rates.

In this report, we compare x264 and LCEVC x264 over a **full encoding ladder** in a live streaming scenario, focusing on gaming and sports, and incorporating the folowing data:

- **Subjective Mean Opinion Score (MOS)** comparisons performed by GB Tech (above left).
- **Objective Quality Metrics: VMAF comparisons** computed with FFmpeg with Rate Distortion Curves and BD-Rate stats computed in Excel (above right).
- **Encoding requirements to** identify the respective encoding costs for the two approaches.
- **Playback power consumption** testing to assess whether LCEVC x264 will drain power faster than hardware accelerated h264 playback.

# Executive Summary

- **The results show that the LCEVC x264 ladder simultaneously saves bandwidth and improves the overall quality of experience:**
  - LCEVC x264 reduced streaming bitrates by ~**18%,** including a 25% bitrate reduction of the top profile. We computed these numbers via a weighted average computation that incorporates the average playback frequency for each profile in the encoding ladder. Simultaneously, LCEVC x264 increased overall quality subjectively by **0.4-0.6 MOS**, with a slight VMAF improvement;
  - LCEVC x264 proved slightly more efficient during transcoding and **should lower transcoding costs** despite producing higher resolution profiles;
  - During playback, LCEVC x264 consumed less power than hardware-accelerated h264 playback.
- Over extensive subjective and objective testing on twenty 1080p sports and eGames videos, **LCEVC x264 outperformed x264 on subjective MOS and VMAF metrics**, with **average BD-Rates respectively of -45.8% and -32.5%** (i.e., in eGames clips, LCEVC x264 achieves same quality as x264 with less than 60% of the x264 bitrate).
- As measured by BD-RATE, LCEVC x264's advantage over x264 **was higher in actual MOS evaluations** than in VMAF scoring.
- **Results were positive in both genres,** with better performance on eGames (BD-Rate-VMAF -33.1%, BD-Rate-MOS -54.2%), and a solid performance on Sports (BD-Rate-VMAF -38.2%, BD-Rate-MOS -44.0%)

# Test Coordinator

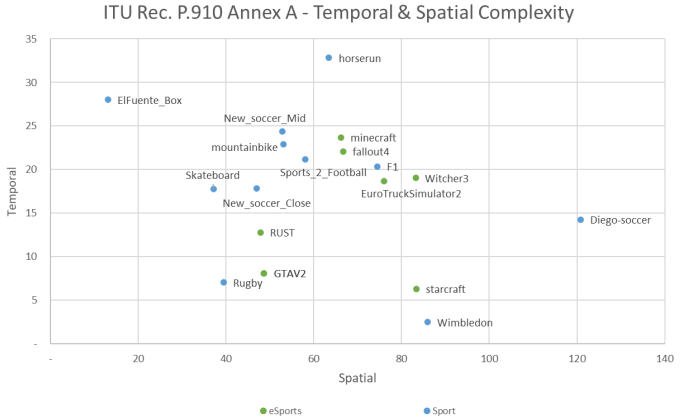
Ozer is a leading expert on H.264, H.265, and VP9 encoding for live and VOD production, the computation and use of video quality metrics, and encoding with FFmpeg. Ozer is a contributing editor to *Streaming Media Magazine*, where he reviews codecs, on-premise and cloud encoders, and ancillary tools like QoE and QoS monitoring services.

Ozer blogs at www.streaminglearningcenter.com, and is the author of over 20 streaming and video-related books, including *Video Encoding by the Numbers: Eliminate the Guesswork from your Streaming Video*, and *Learn to Produce Videos with FFmpeg: In Thirty Minutes or Less*. His books have consistently garnered five-star reviews on Amazon and have been adopted as textbooks by multiple colleges and universities.

# Methodology: Test Clips & Data Rates

**Test clips**

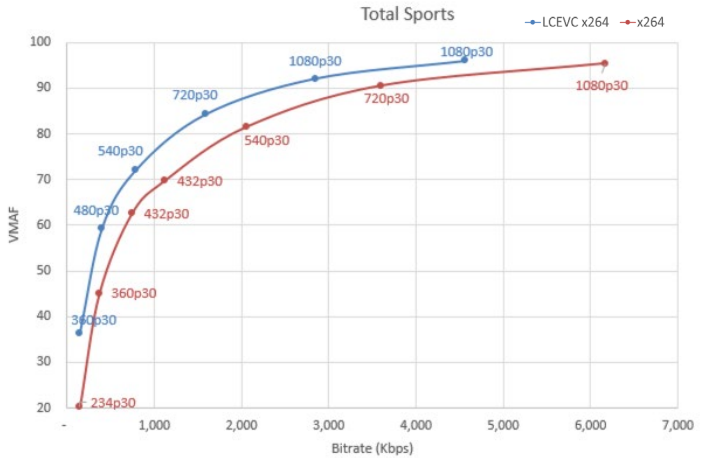
20 test clips from a range of sources in multiple genres including games and sports. As shown on the right, the clips represented an extensive range of temporal and spatial complexity. Clips are presented in Appendix I.




**Data rates and ABR ladder**

We tested at various bitrates (between 145 kbps and 6 Mbps) and resolutions (between 234p and 1080p) to mimic ABR ladders typically used in Live OTT streaming services.

For this reason, the VMAF and MOS scores should be relevant to most use cases.



# Command strings

**x264**

```
ffmpeg -i input.mp4 -r 30 -c:v libx264 -b:v 1000k -bufsize 2000k -g 60.0 -keyint_min 60
-maxrate 1000k -preset medium -sc_threshold 0 output.mp4
```

**LCEVC x264**

```
ffmpeg -i input.mp4 -c:v lcevc_h264 -base_encoder x264 -r 30 -g 60 -b:v 1000k  -
eil_params "preset=medium;scenecut=0;min-keyint=60" output_LCEVC.mp4
```

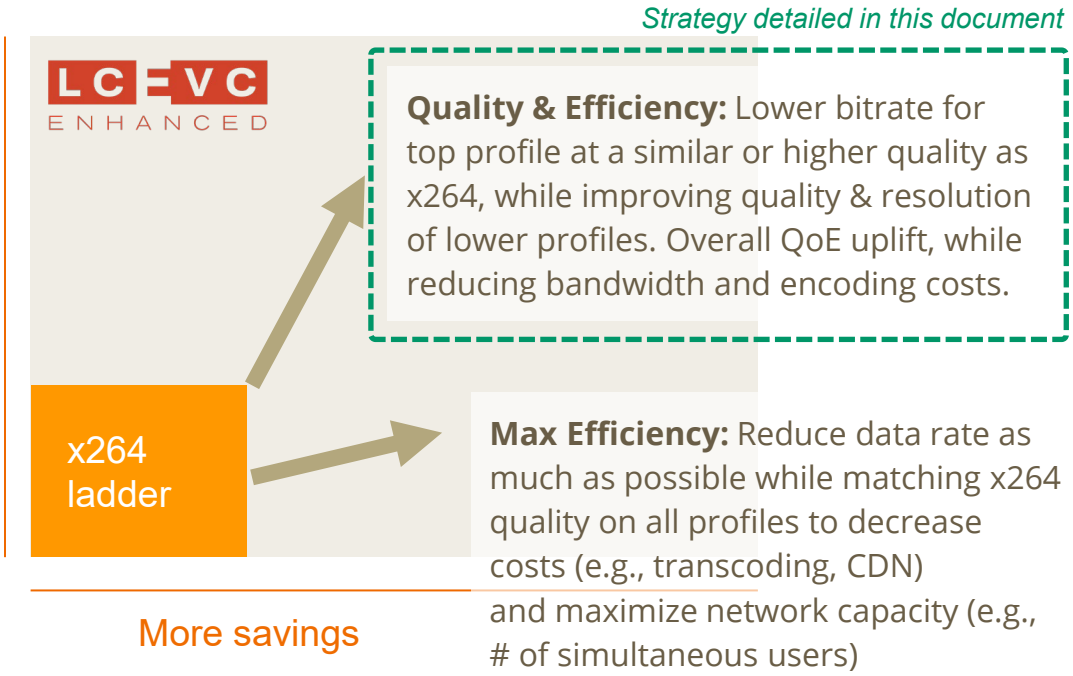
- We tested with a modified build of FFmpeg version 4.3.1 with support for LCEVC enhancement. We used a beta version (v3.1.4-RDSW2, March 2020) of the LCEVC libraries from V-Nova.
- After agreeing on the command strings, V-Nova produced all files with spot verification by Doceo to ensure that the proper script was applied.
- We used **single-pass CBR encoding to simulate the live use case**, with the medium preset for x264 and LCEVC x264 base layer.
- No tuning was used for either codec to achieve maximum visual quality (`lcevc_tune` not specified).
- All sports clips were 1080p30 source while eGames were 1080p60.

# How to build an ABR ladder with LCEVC

**Example x264 Live Sports ABR ladder**

| Profiles | Bitrate (kbps) | Resolution |
|---|---|---|
| #1 | 6,000 | 1080p |
| #2 | 3,500 | 720p |
| #3 | 2,000 | 540p |
| #4 | 1,100 | 432p |
| #5 | 730 | 432p |
| #6 | 365 | 360p |
| #7 | 145 | 234p |

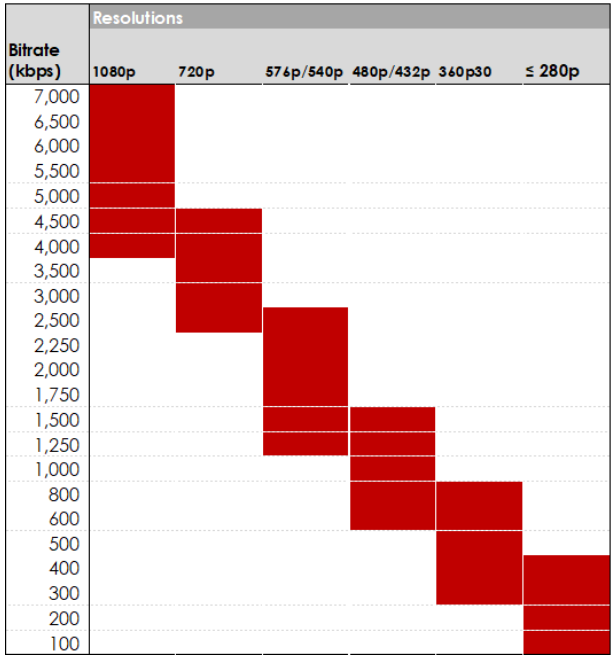
## Different strategies are available to leverage LCEVC in streaming



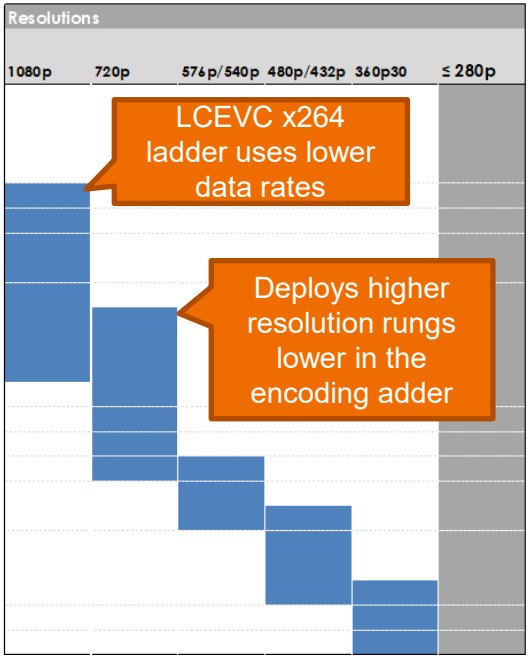
# Optimal resolutions for LCEVC differ from h264

## Recommend streaming resolutions for x264 and LCEVC x264

x264




LCEVC x264




- When converting an existing H.264 ladder to LCEVC x264, producers should use different resolutions and data rates that fully exploit LCEVC
- Specially, at any given bitrate, LCEVC operates more efficiently at a higher resolution than that of the native codec (in this case LCEVC x264 vs. native x264)
- Higher resolutions rungs leverages LCEVC x264's multi-layered approach (base + enhancement) and produces more accurate details and sharpness, while avoiding H.264 artifacts in challenging scenes
- Producers should avoid rungs below 360 in resolution, if necessary dropping the frame rate to meet the data rate target

# Creating the LCEVC x264 Ladder

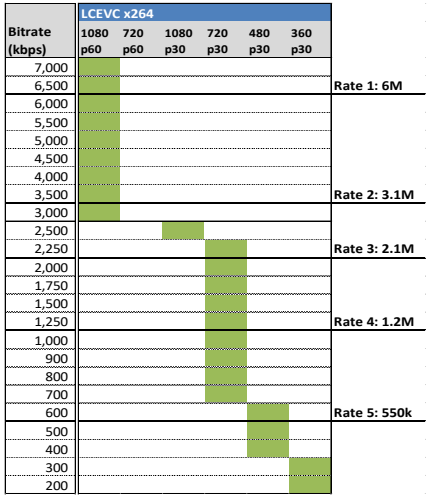
1. To create the LCEVC x264 ladder, we started with a basic encoding ladder for x264 (example on the right).

| Profiles | Bitrate (kbps) | Resolution |
|---|---|---|
| #1 | 6,000 | 1080p |
| #2 | 3,500 | 720p |
| #3 | 2,000 | 540p |
| #4 | 1,100 | 432p |
| #5 | 730 | 432p |
| #6 | 365 | 360p |
| #7 | 145 | 234p |

2. We then tested to find the bitrate where a full resolution LCEVC x264 file matched the VMAF quality score for the top x264 profile.

3. From there, we identified the data rates for the other profiles that would preserve the 1.5x – 2x data rate jump recommended by Apple in TN2224.

4. We identified the optimal LCEVC x264 resolution for each lower profile by encoding at multiple resolutions and choosing the resolution that delivered the highest VMAF quality score, a technique debuted by Netflix (see: http://bit.ly/NF_chull). You see a portion of this analysis on the right.




5. This analysis produced two different LCEVC x264 ABR ladders: one for eGames and one for sports.

# eGames ladder

| x264 | | | | | |
|---|---|---|---|---|---|
| | Bitrate (kbps) | Rez | Fps | Estimated usage | Bitrate/ second |
| Profile 1 | 6,000 | 1080p | 60 | 60.0% | 3,600 |
| Profile 2 | 3,100 | 720p | 60 | 15.0% | 465 |
| Profile 3 | 2,100 | 720p | 30 | 10.0% | 210 |
| Profile 4 | 1,200 | 480p | 30 | 8.0% | 96 |
| Profile 5 | 550 | 360p | 30 | 7.0% | 39 |
| | **12,950** | | | | **4,410** |

| LCEVC x264 | | | | | |
|---|---|---|---|---|---|
| | Bitrate (kbps) | Rez | Fps | Estimated usage | Bitrate/ second |
| Profile 1 | 4,500 | 1080p | 60 | 67.5% | 3,038 |
| Profile 2 | 2,700 | 1080p | 60 | 12.5% | 338 |
| Profile 3 | 1,500 | 720p | 30 | 8.5% | 128 |
| Profile 4 | 800 | 540p | 30 | 6.0% | 48 |
| Profile 5 | 450 | 480p | 30 | 5.5% | 25 |
| | **9,950** | | | | **3,575** |

-25%

-23%

-19%

*LCEVC x264 saving on total bitrate*

*LCEVC x264 saving on average streamed bitrate*

To compute the bandwidth savings over the entire ladder we estimated the **average usage for each profile** after conferring with several OTT providers. This estimated usage allowed us to compute the estimated per-second bandwidth consumption for each technology, leading to the savings identified below.

These are **"weighted-average" savings**, which more closely approximate the savings achievable in real-world deployments than traditional BD-Rate computations because they consider the estimated usage of the different profiles. In contrast, BD-Rate computations assume that all profiles are consumed equally, which seldom happens in real-world deployments.

**The LCEVC x264 ladder outperformed x264 as follows:**

- **25% bitrate reduction on top profile**, while delivering higher quality vs. the same profile of the x264 ladder (higher MOS, similar or greater VMAF).
- **19% reduction on average streamed bitrate** (e.g., CDN costs), 23% on total bitrate (e.g., storage and egress costs).

# Sports ladder

| x264 | | | | | |
|---|---|---|---|---|---|
| | Bitrate (kbps) | Rez | Fps | Estimated usage | Bitrate/ second |
| Profile 1 | 6,000 | 1080p | 30 | 71.6% | 4,296 |
| Profile 2 | 3,500 | 720p | 30 | 13.5% | 473 |
| Profile 3 | 2,000 | 540p | 30 | 9.5% | 190 |
| Profile 4 | 1,100 | 432p | 30 | 3.2% | 35 |
| Profile 5 | 730 | 432p | 30 | 1.2% | 9 |
| Profile 6 | 365 | 360p | 30 | 0.6% | 2 |
| Profile 7 | 145 | 234p | 30 | 0.4% | 1 |
| | **13,840** | | | | **5,005** |

| LCEVC x264 | | | | | |
|---|---|---|---|---|---|
| | Bitrate (kbps) | Rez | Fps | Estimated usage | Bitrate/ second |
| Profile 1 | 4,500 | 1080p | 30 | 77.0% | 3,465 |
| Profile 2 | 2,700 | 1080p | 30 | 11.0% | 297 |
| Profile 3 | 1,500 | 720p | 30 | 8.0% | 120 |
| Profile 4 | 800 | 540p | 30 | 2.5% | 20 |
| Profile 5 | 400 | 480p | 30 | 1.0% | 4 |
| Profile 6 | 145 | 360p | 30 | 0.5% | 1 |
| Profile 7 | | | | | |
| | **10,045** | | | | **3,907** |

-25%

-27%

***LCEVC x264 saving on total bitrate***

-22%

***LCEVC x264 saving on average streamed bitrate***

We adopted the x264 sports ladder from the encoding ladder recommended by Apple in the HLS Authoring Specification (apple.co/2RumLOZ). Estimate usage was derived from the middle ladder shown in this Brightcove White Paper (http://bit.ly/bc_contextaware, see Table 3, additional data supplied by Brightcove)

**The LCEVC x264 ladder outperformed x264 as follows:**

- **25% bitrate reduction on top profile**, while delivering higher quality vs. the same profile of the x264 ladder (higher MOS, similar or greater VMAF)
- **22% reduction on average streamed bitrate** (e.g., CDN costs), 27% on total bitrate (e.g., storage and egress costs)

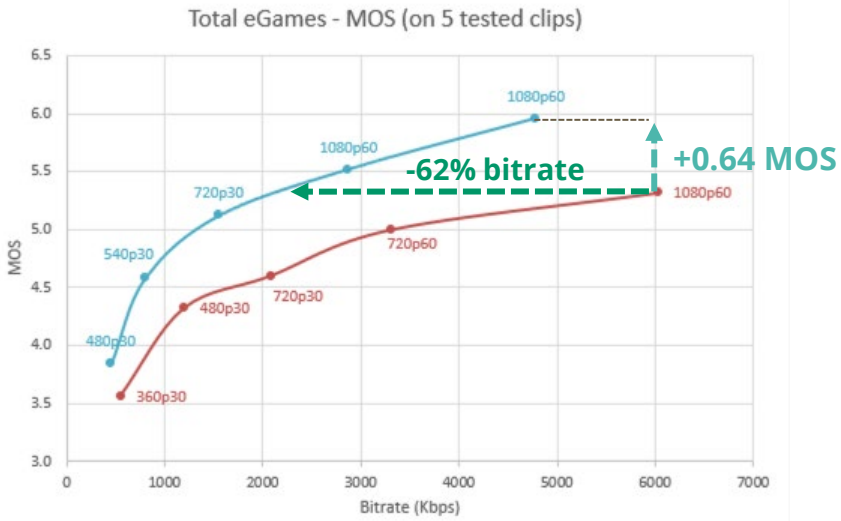
# Subjective Quality Testing

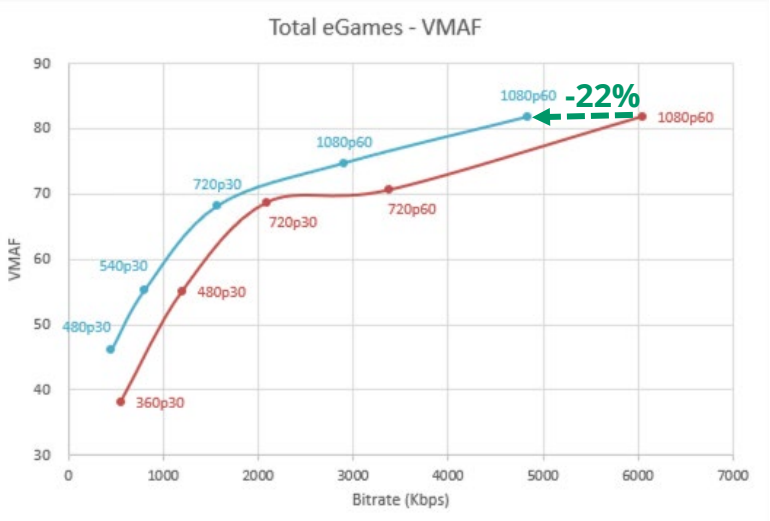
- After creating the eGames and Sports ladders, we encoded all tests files as specified above.
- We computed VMAF using FFmpeg on all profiles in all ladders for both x264 and LCEVC x264. These computations provided the VMAF scores we'll show in the upcoming slides.
- Subjective comparisons were performed by GBTech using the Double Stimulus Impairment Scale (DSIS) protocol, which operates as follows:
  - The viewer is shown the source video.
  - Then the encoded video.
  - Then they score the clip using the scale shown on the right.
  - This provided the Mean Opinion Scores shown on the following slides.

## DSIS impairment scale

| Score | Impairment item | |
|---|---|---|
| 10 | Imperceptible | |
| 9 | Slightly perceptible | somewhere |
| 8 | | everywhere |
| 7 | Perceptible | somewhere |
| 6 | | everywhere |
| 5 | Clearly perceptible | somewhere |
| 4 | | everywhere |
| 3 | Annoying | somewhere |
| 2 | | everywhere |
| 1 | Severely annoying | somewhere |
| 0 | | everywhere |

# eGames: MOS & VMAF Rate distortion curves



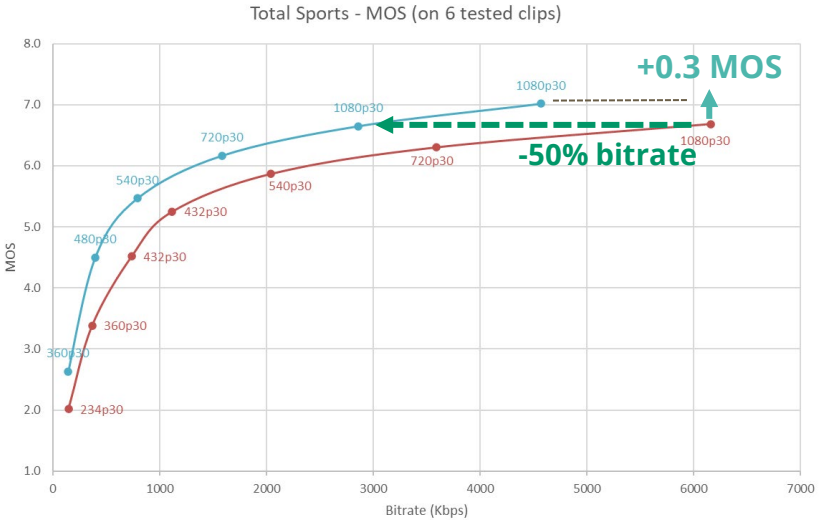

The subjective comparisons showed that we were too conservative when choosing our top rate based on VMAF, with LCEVC x264 delivering the same quality as x264's 1080p60 clip at about a **62% bitrate reduction**.

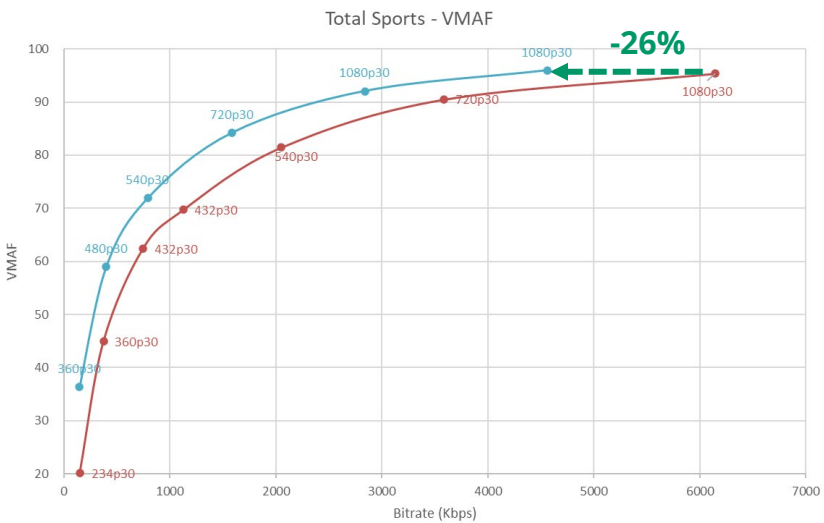
Focusing on the top profile, even after **reducing the bitrate by 25%**, LCEVC x264 delivered **0.64 additional MOS points** when comparing profile 1 encodes (1080p60).

With VMAF, our top rate estimates were more accurate, with LCEVC x264 delivering the same quality in the top profile with a bitrate reduction of about 22%.

Both MOS and VMAF results prove that LCEVC x264 delivers **better quality throughout the encoding ladder**, improving QoE and reducing churn, **while reducing bandwidth consumption by ~15%.**

# Sports MOS and VMAF Rate distortion curves





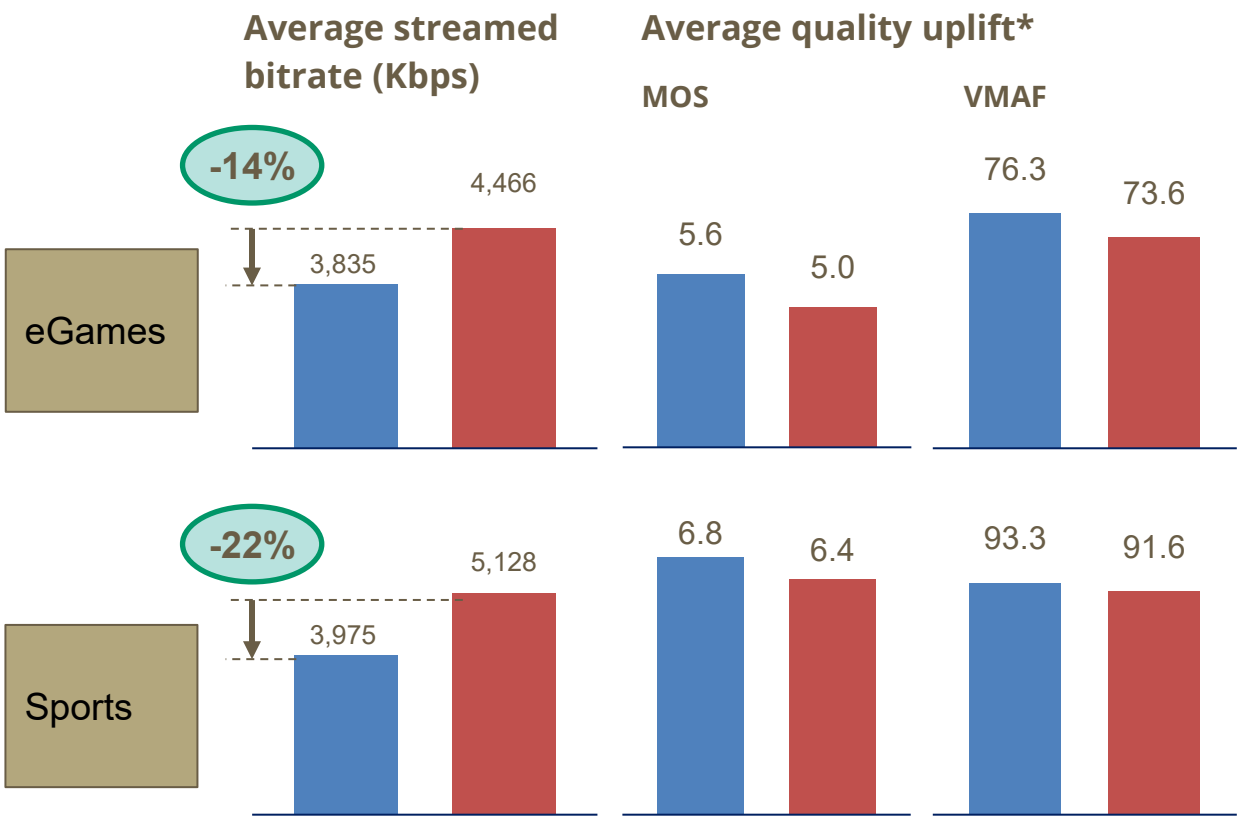
The sport clips comparisons showed similar results. Specifically, we could have **dropped the top LCEVC x264 profile by about 50%** and achieved the same quality as the top x264 profile.

Again, even after **reducing the bitrate by 25%,** LCEVC x264 delivered **.3 additional MOS points** when comparing profile 1 encodes (1080p30).

VMAF results confirmed our goal of retaining about the same quality as x264 in the top profile while reducing the bitrate by 26%.

And again, both MOS and VMAF results confirm that LCEVC x264 delivers better quality in all rungs of the encoding ladder, improving QoE and reducing churn, while reducing bandwidth consumption by ~22%.

# Key results: 20% bitrate savings + higher quality




- We computed the bitrate savings in a previous slide, which showed a **14-22% reduction of** average streamed bitrates, which will **reduce CDN streaming costs**
- We computed the quality impact on the respective encoding ladders using the same procedure shown for bitrate savings; substituting VMAF/MOS values for data rate and using the same usage statistics to compute overall MOS and VMAF for the x264 and LCEVC x264 ladders.
- These results showed that LCEVC x264 **improved quality by .6 MOS for eGames and .4 MOS for Sports, and 2.7/1.7 VMAF points respectively.**



* Weighted average of MOS/VMAF taking into account the frequency of viewing of each profile

# Per content quality improvement

Subjectively tested contents

## eGames ABR ladder

| # | Content | Weighted average MOS | | Weighted average VMAF | |
|---|---|---|---|---|---|
| | | x264 | LCEVC x264 | x264 | LCEVC x264 |
| 1 | EuroTruckSim2 | 5.0 | 5.3 | 73.4 | 75.1 |
| 2 | fallout4 | - | - | 77.7 | 79.6 |
| 3 | GTAV | - | - | 75.5 | 78.5 |
| 4 | GTAV2 | 6.0 | 6.3 | 91.8 | 95.5 |
| 5 | minecraft | 4.1 | 5.5 | 64.0 | 66.1 |
| 6 | RUST | - | - | 66.5 | 69.4 |
| 7 | starcraft | 6.2 | 7.0 | 87.4 | 93.5 |
| 8 | WITCHER3 | 3.6 | 4.0 | 52.7 | 52.6 |
| **Total eGames** | | **5.0** | **5.6** | **73.6** | **76.3** |
| Total Subjectively tested | | **5.0** | **5.6** | **73.9** | **76.6** |

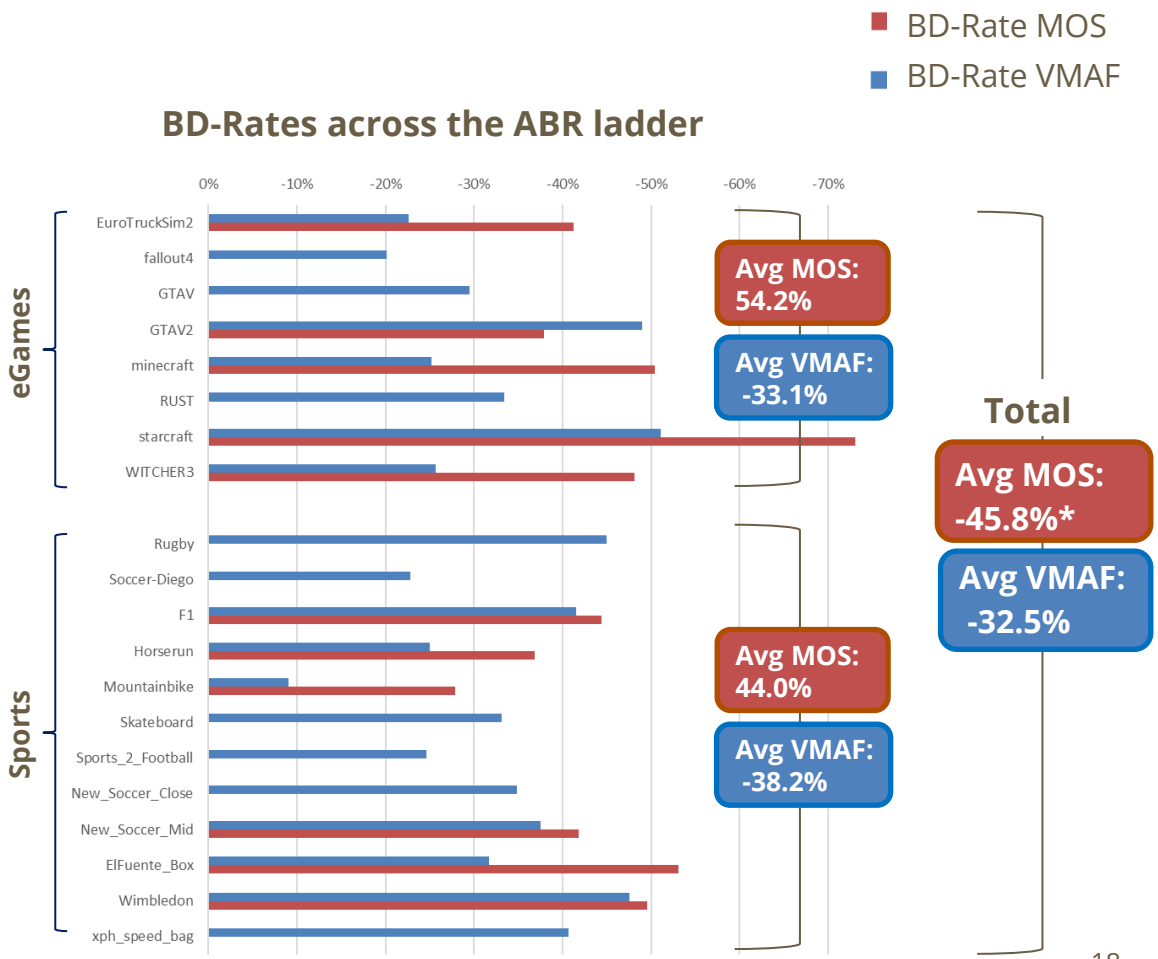
## Sports ABR ladder

| # | Content | Weighted average MOS | | Weighted average VMAF | |
|---|---|---|---|---|---|
| | | x264 | LCEVC x264 | x264 | LCEVC x264 |
| 1 | Rugby | - | - | 86.3 | 90.5 |
| 2 | Soccer-Diego | - | - | 84.6 | 84.0 |
| 3 | F1 | 6.2 | 6.7 | 89.7 | 92.6 |
| 4 | Horserun | 5.9 | 6.3 | 84.4 | 83.9 |
| 5 | Mountainbike | 6.7 | 6.7 | 95.3 | 93.5 |
| 6 | Skateboard | - | - | 93.5 | 93.9 |
| 7 | Sports_2_Football | - | - | 91.1 | 88.9 |
| 8 | New_Soccer_Close | - | - | 94.1 | 96.0 |
| 9 | New_Soccer_Mid | 6.2 | 6.5 | 93.7 | 95.8 |
| 10 | ElFuente_Box | 7.0 | 7.5 | 97.6 | 97.9 |
| 11 | Wimbledon | 6.7 | 7.4 | 92.3 | 97.9 |
| 12 | xph_speed_bag | - | - | 96.1 | 97.5 |
| **Total Sports** | | **6.4** | **6.8** | **91.6** | **92.7** |
| Total Subjectively tested | | **6.4** | **6.8** | **92.2** | **93.6** |

- Here the raw score for each tested clip
- The weighted average is calculated according to the estimated average usage for each profile

# Per content BD-rates

- We also computed traditional BD-rates across all points in the ABR ladder. Despite being a ‘theoretical’ measure in the context of the ladder (they assume same playback frequency for all profiles), they are a widely adopted method to compare codecs.
- Subjective MOS BD-Rate scores (red) and VMAF scores (blue) show that when encoding the ABR ladder, ***LCEVC x264 is able to produce the same visual quality as x264 at about between 54.2% (MOS) to 67.5% (VMAF) of the data rate.***



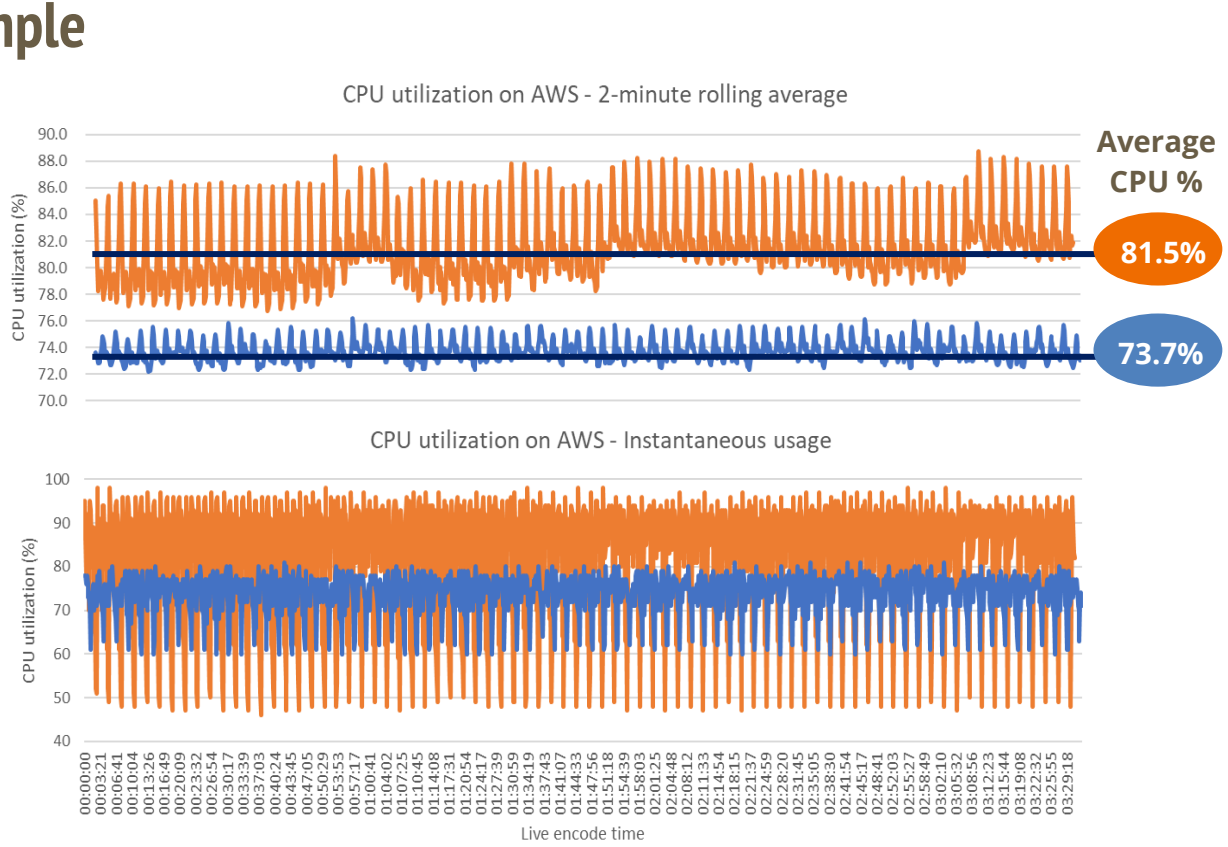

* MOS is calculated on a subset of 11 clips. The corresponding BD-Rate-VMAF for the same 11 clips is -33.3%

# Encoding Requirements for the full ABR ladder

## eGames Example

**Methodology**

- Both ladders encoded in the same AWS c5a.8xlarge instance on V-Nova's cloud encoding platform
- LCEVC x264 ABR ladder included: 2x 1080p60 profiles, 1x 720p30, 1x 540p30,1x 480p30
- x264 ladder included: 1x 1080p60 profile, 1x 720p60, 1x 720p30, 1x 480p30, 1x 360p30




- **LCEVC x264 consumed 11% less CPU (or 7.8 p.p. difference) than x264** despite 1.3x more encoded pixels and higher quality
- The LCEVC ladder reliably fits the AWS instance, with **CPU load consistently <75%**, while h.264 oscillates close to saturation levels

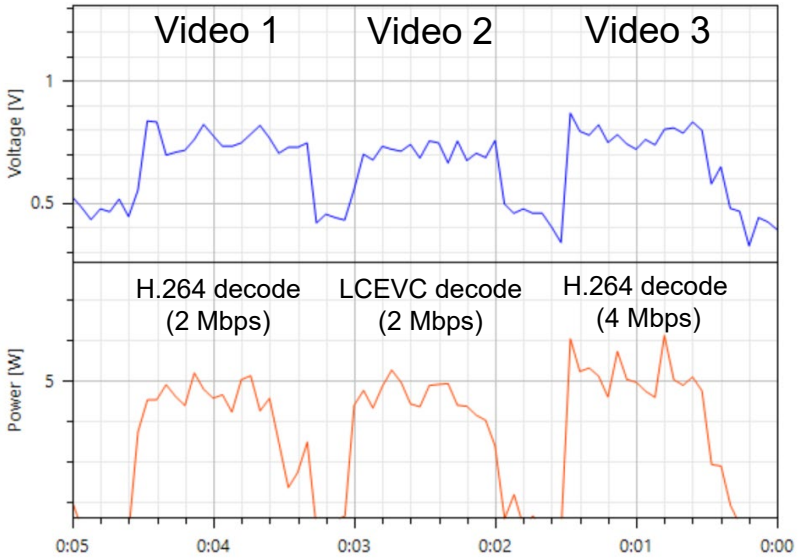
# Decoder power drain

Most mobile platforms use hardware-accelerated H.264 playback to reduce battery consumption while playing H.264-encoded videos. With LCEVC, decoding the H.264 base-layer is also hardware accelerated, though decoding the enhancement layer is not.

We used the following tests to compare H.264 and LCEVC battery consumption. The tests involved three video files.

- Video 1 – H.264 @ 2 Mbps
- Video 2 – LCEVC @ 2 Mbps
- Video 3 – H.264 @ 4 Mbps (to match LCEVC quality)

We played these files back on a Zotac Zbox-EN72080v computer with a six-core I7-9750H running Windows 10, measuring voltage and power consumed with the Open Hardware Monitor utility (https://openhardwaremonitor.org/).




As you can see, compared to the 2 Mbps H.264 file, **LCEVC decode consumes lower voltage and about the same power**, so overall, LCEVC decode consumes less battery power. Compared to the 4 Mbps H.264 file, which is the **same approximate quality** as the LCEVC file, **LCEVC playback is more efficient in both power and voltage**.

So, **despite the lack of hardware acceleration for the decode of the enhancement layer, LCEVC playback is slightly more efficient than H.264 playback**.

# Work to Come – Stay tuned

- These findings relate solely to LCEVC used to enhance h264, with resolution up to 1080p
- Further research will explore LCEVC's performance when enhancing other codecs (e.g., x265, AV1), as well as UHD resolution