360 8K Viewport-Independent VR

Thierry Fautier

Harmonic, USA

Patrick Gendron

Harmonic, France

Xavier Ducloux

Harmonic, France

Vincent Lepec

Viaccess-Orca, France

Pascal Perrot

Orange, France

Written for presentation at the

SMPTE 2021 Annual Technical Conference & Exhibition

Abstract. *360 VR has been deployed in the past few years using different techniques. Viewport-independent technology is used on 4K content for delivery to 4K-capable head-mounted displays (HMDs) and smartphone devices, resulting in a disappointing experience. The alternative is using viewport-dependent technology with 8K content on 4K-capable HMDs and smartphones devices, which enables a good experience, but with complexities and limitations in terms of the integration into existing OTT workflows. The 8K viewport-independent technology presented uses off-the-shelf encoding techniques to compress 8K 360 VR content as a single file and to distribute it in CMAF low-latency DASH mode to 8K-capable devices such as the Oculus Quest 2 or Galaxy S20. This paper will present an end-to-end 8K VR workflow, which is entirely cloud based and capable of delivering high-quality 8K VR DRM-protected content compared with 4K content on different devices. The paper will present the initial results of the trial performed by the VR Study Group of the Streaming Video Alliance (SVA) in collaboration with the VR Industry Forum, where 8K VR content was encoded and streamed on different very-high-speed (fiber and DOCSIS) networks.*

Introduction

VR is a technology that has been developed over the years and evolved. Figure 1 describes the generic VR workflow. The content is acquired with a volumetric 3D camera system, optionally stitched, and then mapped to a set of different projections (equirectangular, cubic). Afterward, content is encoded either in a viewport-independent (the entire sphere is sent) or viewport-dependent (only watched region is transmitted) way. The content is encoded and transmitted using either single bitrate for viewport dependent or adaptive bitrate for viewport independent. The decoding can happen either on a flat device (like PC or mobile), or on an HMD where the experience is fully immersive. [1] provides more details of the different stages of processing.

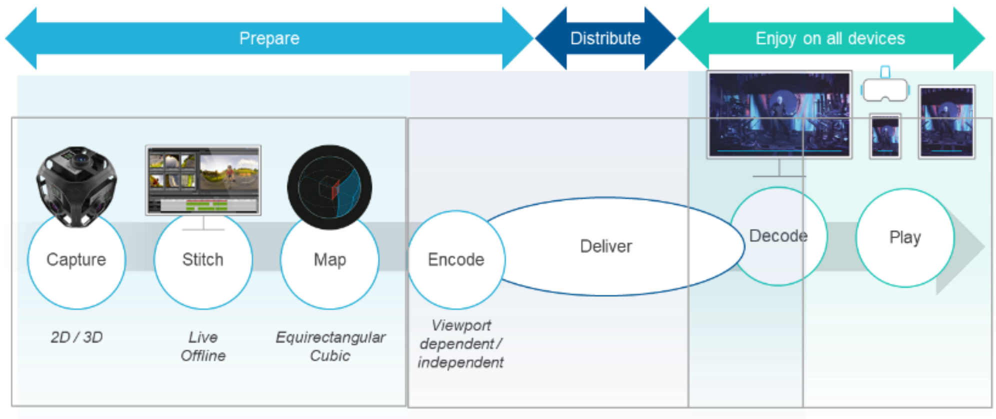

Figure 1 - Generic VR workflow

VR has started with a 4K viewport-independent experience where a 4K capture system was used, and then content was stitched, encoded and sent over the internet to a 4K-capable phone connected to HMDs. The experience was not satisfactory at that time, which we believe has hampered the take-off of VR.

The second generation came with a viewport-dependent technology where an 8K capture was sliced in tiles that were only transmitted if they corresponded to the field of view of the user's device. [1] describes a 4K viewport-dependent solution. This approach offers two main advantages: the capability to work with 4K-capable devices and the ability to reduce the bandwidth of the transmitted video, as only the consumed content was transmitted. This technology has several drawbacks: it requires complex integration on the encoding, encryption, streaming and player side; when tiles do not arrive on time they will provide a fuzzy video for a duration of several frames; it is not integrated into an existing (over the top) OTT workflow, as different protocols are used; and the encoder and decoder are provided by the same company, limiting the choices for operators.

The third wave of VR arrived when 8K devices were released in 2020 with 5G mobile phones and HMDs, such as Oculus Quest 2, Qualcomm XR2, the Skyworth 901 and Pico 3. All of these devices have a built-in 8K decoder. If the resolution of capture goes beyond 8K (i.e., 12K, 16K), then the viewport-dependent technology could find a new application on 8K-capable devices, providing the "ultimate experience." Figure 2 summarizes the different VR phases.

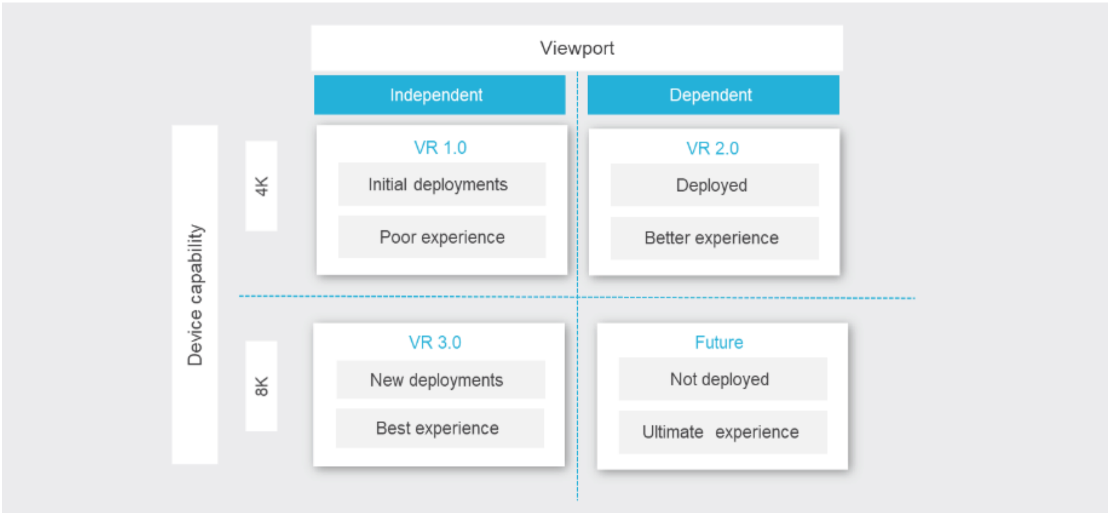

Figure 2 – VR Phases

This paper will present the details of an end-to-end, standard-based 8K viewport-independent solution and the results of the first trial conducted inside the Streaming Video Alliance (SVA) VR Study Group.

8K Viewport-Independent Concept

VR has evolved from initially being 4K viewport independent to 4K viewport dependent to double the resolution per eye in the viewport and providing a better quality of experience. With the advent of new 8K-capable devices — including Qualcomm XR2 and next-generation 5G smartphones released in 2020 — it is now possible to send and decode the full 8K sphere in the device and provide the same perceptible resolution in the viewport as the best 4K viewport-dependent solutions. The viewport-independent solution has several benefits:

- The VR application solution can be hooked to an existing OTT workflow, as the same DASH protocol is used. It also uses DASH-associated OTT benefits: (Common ENCryption) CENC, (Digital Rights Management) DRM, (Common Media Application Framework) CMAF low latency, digital ad insertion and use of existing OTT analytics. Therefore, this solution enables service continuity, including the capability to mix 2D and VR 360 both transported in DASH.
- There is no delay in receiving the content. At worst, the player will pick a lower profile in adverse network conditions.
- The encoder is completely independent of the decoder, leaving more choice for operators.

The proposed viewport-independent approach fits quite well with the deployment of very high-speed broadband networks such as fiber, DOCSIS 3.1 and 5G, where gigabit network connectivity is now getting within reach. As the whole sphere is being transmitted vs. a viewport-dependent approach, more storage and bandwidth are required on the network. However, we only expect a 50% increase vs. 4K viewport dependent, when using state-of-the-art, cloud-based Content Aware Encoding (CAE) technology featuring HEVC compression.

Table 1 shows the difference between viewport-dependent and viewport-independent technologies.

Features	4K viewport independent	4K viewport dependent	8K viewport independent
Standard based	OMAF 1.0	OMAF v1.0/OMAF v2.0	Extension of OMAF 2.0
Quality	No lag in any network conditions Limited (HD) perceived resolution	Lag in stressed network conditions Limited ABR support	No lag in any network conditions Reduced resolution on 4G
Decoder performance	4K decoder required	4K+ decoder required	8K decoder required
Decoder base	All head-mounted displays (HMDs)	All high-performance 4K HMDs/phones	Gear VR/S10 Skyworth v901 Qualcomm XR2 Oculus Quest 2 S10/S20 phones
Protocol	OMAF 1.0/DASH	OMAF 2.0	OMAF 2.0/DASH
Latency	5-7s (1)	20-30s	5-7s (1)
DRM integration	Standard	Non-standard/complex	Standard
Multi-client integration	Standard	Non-standard/complex	Standard
DAI	Standard	Non-standard/complex	Standard
Bitrate (top profile)	10-15 Mbps	20 Mbps	35 Mbps
Analytics	Use OTT analytics	Develop custom analytics	Use OTT analytics
CDN independent	Yes	No (2)	Yes
Visual experience	Poor	Good	Good

Table 1 – Technology comparison

(1) When DASH CMAF low latency is used

(2) Requires CDN optimization for best performance

VRIF Adoption

VR-Industry Forum (VR-IF) has developed a new distribution profile for viewport-independent delivery of 8K video content in its Guidelines version 2.2 (2) that will match the same resolution per eye as achieved by the 4K viewport-dependent (Omnidirectional Media Application Format) OMAF profile today. The 8K viewport-independent profile inherits all of the properties of the OMAF viewport-independent profile, but it requires support for HEVC Main 10 Level 6.1

decoding capability in order to process 8K content. VR-IF believes that this is a pragmatic means to increase the adoption of high-quality live VR experiences by leveraging 8K processing capabilities of new 5G mobile devices, including phones and next-generation HMDs, such as the Oculus Quest 2 released at the end of 2020.

DASH

DASH is now becoming universal for streaming services, and we believe that in order to offer a compelling VR service and to leverage existing infrastructure deployed for OTT services the service has to be based on DASH. The legacy DASH framework enables low latency with CMAF, seamless DRM integration with CENC and support for digital ad insertion and content replacement.

Encoder

Using CAE, content providers can deliver VR at the lowest bitrate and at the highest quality. The fundamentals of CAE, as implemented by Harmonic with its AI-based EyeQ® technology, are described in detail (3). The Ultra HD Forum recommends using CAE for 4K encoding and, of course, the same technique can also be used for 8K. Some early results of 8K encoding using CAE for streaming applications have been reported (4).

Table 2 describes the encoding profiles being used. As CAE is used, an encoding cap is defined as well as the average encoded rate, which depends entirely on the complexity of the content.

Profiles	Cap (Mbps)	Resolution	Frame rate (fps)	HEVC encoded bitrate (Mbps)
8K	42 Mbps	7680x4320	30 fps	30-35 Mbps
4K high quality (HQ)	25 Mbps	3840x2160	30 fps	15-18 Mbps
4K standard quality (SQ)	15 Mbps	3840x2160	30 fps	10-12 Mbps
1080p HQ	8 Mbps	1920x1080	30 fps	3.5-5 Mbps
1080p SQ	5 Mbps	1920x1080	30 fps	2.5-3.5 Mbps
720p	3 Mbps	1280x720	30 fps	1.5-2.5 Mbps

Table 2 – 8K encoding profiles

Decoder Capabilities

The following devices have been tested using 8Kp30/60 content monoscopic/stereo, encoded in HEVC Main 10 and packaged with DASH. Table 3 presents the progress on interoperability.

Device	Video format	Chipset	Player	Mode
Galaxy S10	8Kp30	SDM 855	Viaccess-Orca	Magic window
Gear VR/Galaxy S10	8Kp30	SDM 855	Viaccess-Orca	HMD
Skyworth v901	8Kp30	Exynos *8895*	Native player	HMD
Qualcomm XR2	8Kp60	SDM 865	Viaccess-Orca	HMD
Oculus Quest 2	8Kp60	SDM 865	Viaccess-Orca	HMD
Galaxy S20	8Kp60	SDM 865	Viaccess-Orca	Magic window

Table 3 - Interoperability table

Backward Compatibility

Backward compatibility with 4K-only devices, such as the Oculus Quest, is ensured by using 4K profiles and below. Moreover, when delivered through a limited capacity network like 4K, even if the device is 8K capable, it might have to fall back on 4K and below profiles.

Figure 3 describes the match between content and devices in the different scenarios. Content is encoded in ABR from 8K, down to 4K and HD. The 5G network is capable of streaming any type of resolutions, and as most of 5G devices can decode 8K, they can process the 8K stream. On a 5G network, when the phone is the modem, 4K only capable devices will pick the 4K profile. On a 4G network, devices will only pick 4K resolution, due to the bandwidth limitation.

In conclusion, an 8K ABR scheme can adapt to all kinds of network and device combinations.

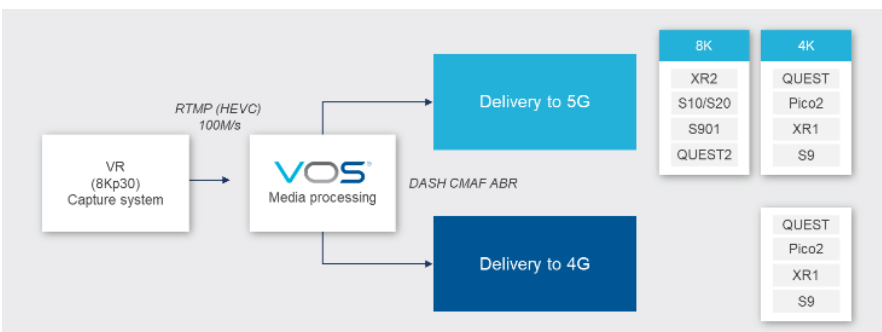

Figure 3 – Different scenarios for 8K VR ABR delivery

Bitrates

Table 4 - Formula 3 8K encoding profiles for VR IF test clips of Formula 3 content encoded with an 8K profile ladder, highlights the different bitrates.

Profiles	Resolution	Frame rate (fps)	Average video bitrate (Mbps)	Max video bitrate measured over chunk duration (Mbps)
8K Cap 42 Mbps	7680x4320	25 fps	28.6-32 Mbps	42.6 Mbps
4K Cap 25 Mbps	3840x2160	25 fps	16.6-18.3 Mbps	26 Mbps
4K Cap 15 Mbps	3840x2160	25 fps	10.2-12.1 Mbps	15.5 Mbps
1080p Cap 8 Mbps	1920x1080	25 fps	3.6-5 Mbps	6.3 Mbps
1080p Cap 5 Mbps	1920x1080	25 fps	2.2-3.6 Mbps	4.9 Mbps
720p Cap 3 Mbps	1280x720	25 fps	1.3-2.4 Mbps	2.9 Mbps

Table 4 - Formula 3 8K encoding profiles

The 8K profile, compared with a high-quality 4K profile (with a cap of 25 Mbps), is only 74% higher. As CAE compression is used, we also indicate the top bitrate measured during a chunk duration, and the 8K bitrate tops out at 42.6 Mbps versus 26 Mbps for 4K, a 62% increase.

It is important to note that these data are provided for live encoding with a prototype encoder, meaning that they are not optimized yet. Over time, it's expected that 8K will not be more than 50% higher than 4K viewport dependent.

Low Latency

Using standard DASH protocol brings all the associated features to DASH such as low latency using CMAF Low Latency Chunk Transfer Encoding, and we can achieve a delay of five to seven seconds from capture to rendering, as detailed in (5). This enables a delay that is close to the broadcast delay and therefore enables to have a social interaction with others, including being able to exchange instant messages in conjunction with the VR experience. This functionality is not available when using other protocols that do not have a built-in low latency mode, such as such as viewport dependent.

Power

Power consumption is always an important factor to consider when moving to a new technology, not only on the network side but also on the device side, as it is widely recognized that the carbon footprint of a video service is largely based on the device consumption (6). We have, therefore, decided to measure the power impact when moving to more recent HMDs, such as the Quest 2, and see the impact of moving from 4K to 8K.

Regarding device autonomy, the 8K viewport-independent playback leads to 40% to 50% more power consumption than 4K playback, measured on a Galaxy S10 phone.

Gear VR, using the Galaxy S10, shows seven hours of autonomy on 4K content, while with 8K, it is reduced to four hours (reduction of 40%). Here we can see the HMD vs. flat usage does not change the power dissipation.

We have conducted a comparison on the Quest 2, using a XR2 chipset that is one generation ahead of the chip used in Gear VR/S10. We measured the difference between 4K and 8K content as only 5% less autonomy (out of three hours autonomy), meaning we now have nearly no power consumption difference between 4K and 8K compared with the previous generation of HMDs.

Tables 5 and 6 summarize the different measurements made and have normalized the autonomy of Gear VR/S10 to a 100 scale as a reference.

Source	**Device**		
	Galaxy S10	Gear VR	Chipset
4K	100	100	SDM 855
8K	60	60	

Table 5 – Resolution impact on first generation HMD/phone autonomy

Source	**Device**		
	Galaxy S20	Quest	Chipset
4K	NA	58	SDM 865
8K	NA	55	

Table 6 - Resolution impact on second-generation HMD/phone autonomy

It is interesting to note that the Quest 2 (independent of the input resolution) has a much lower autonomy than the Gear VR (no comparison with Quest 1 available), due to the fact that the device has much more functionality available.

It is important to note that moving from 4K to 8K increases the number of pixels processed by 300% and the bitrate by 100% while on the Quest 2 we see nearly no power consumption impact.

Secure Player

Description

Using standard DASH protocol brings all the associated features to DASH such as CENC. The player used to decode the 8K viewport-independent encrypted streams is the Viaccess-Orca Secure Player (VO Player) that can decode a DASH CENC encrypted stream using Widevine L1 encryption and render it on an Oculus Quest 2. It provides a unified DASH Ultra Low Latency experience (with five seconds from capture to rendering) relying on a common implementation supported across multiple platforms and ecosystems, as described in Table 7.

Devices	Usage	Mode	Platform	Package provided	DRMs
STB	Stand-alone	Magic window	Android	VO Player SDK	Widevine (L1 & L3), PlayReady
Smart TVs	Stand-alone	Magic window	Android	VO Player SDK	Widevine (L1 or L3), PlayReady
PC	Stand-alone	Magic window	Web	VO Player HTML5 Player	Wide vine (L3), PlayReady
Tablet	Stand-alone	Magic window	Android	VO Player SDK	Widevine (L1 or L3), PlayReady
Mobile	Stand-alone	Magic window	Android / iOS	VO Player SDK	Widevine (L1 or L3), PlayReady
	Plug to HMD (deprecated)	VR	Android	VO Player Unity SDK	Widevine (L3), PlayReady
HMD	PC tethered	VR	WebXR	VO Player HTML5 Player	Widevine (L3), PlayReady
	Stand-alone	VR	Android – Oculus	VO Player Unity SDK	Widevine (L1 & L3), PlayReady

Table 7 – Device platform support

Since the initial VR deployment phases (e.g., VR1.0), the VO Secure Player integrates viewport-independent 360 support in addition to the core OTT functionalities, following the fast evolutions of the VR market (e.g., 8K-capable VR devices).

The player is able to display monoscopic and stereoscopic content, either through a VR-compatible device for stereo, or by displaying only one eye for stereoscopic content on a flat screen. The player can also feature high-order audio ambisonic for immersive audio and video

360 use cases as well as broad OTT analytics to enable QoE and QoS monitoring. In addition, the VO Secure Player features a unique multi-video and camera experience, enabling end users to switch seamlessly between all VR camera available at a sports event.

Game Engines: The Game Changers

In order to address complex interactive project developments involving a transversal technological knowledge (i.e., CGIs, UI/UX, 3D, AR, VR, XR, etc.) the VO Secure Player SDK can be easily integrated within the Unity Game engine.

Unity is a cross-platform game engine developed by Unity Technologies, first announced and released in June 2005 at Apple's Worldwide Developers Conference as a Mac OS X-exclusive game engine. As of 2018, the engine had been extended to support more than 25 platforms. The engine can be used to create three-dimensional, two-dimensional, virtual reality and augmented reality games and applications, as well as simulations and other experiences. The engine has been adopted by industries outside video gaming, such as film, automotive, architecture, engineering and construction.

The following architecture illustrates how the VO Secure Player is integrated in a Unity- and Oculus-based platform (e.g., Oculus Quest v1 and v2 and Qualcomm XR2).

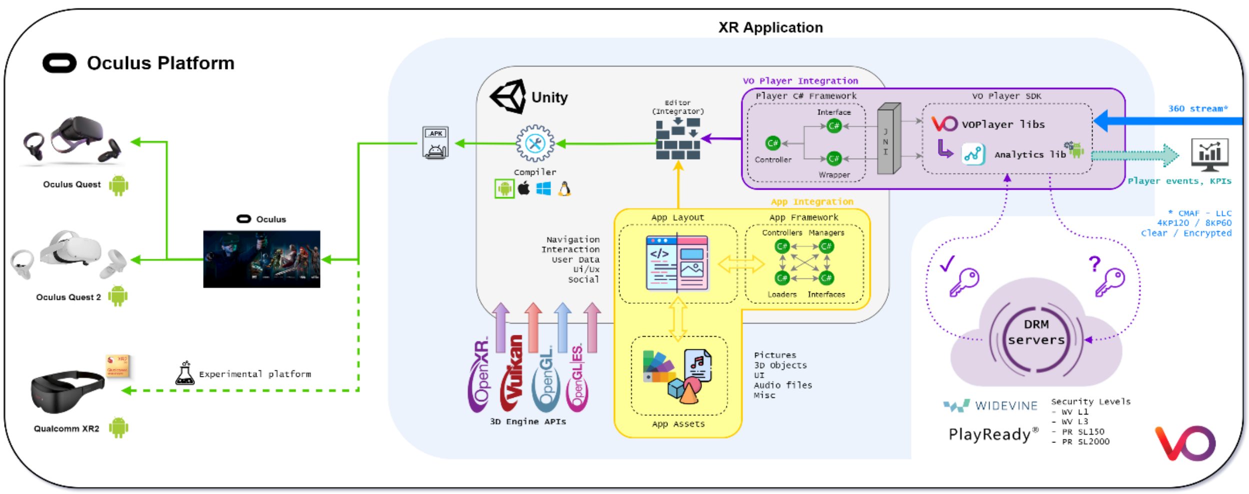

Figure 3 - VO Player VR platform architecture

This novel integration scheme brings convergence to industry standards between the traditional OTT workflows and the next generation of live video experiences (i.e., live sports events, live concerts and immersive documentaries, metaverse), embedding off-the-shelf OTT features (e.g., CMAF compliance, low-latency support, smooth streaming, digital ad insertion, DRMs, multiview and watch together) seamlessly portrayed in those new ecosystems.

Steaming Video Alliance (SVA) Trial

The SVA VR Study Group tested the 8K viewport-independent solution on various very-high-speed networks using a standard CDN.

The next phase of the trial will use open caching deployed on the ISP's fiber network.

Participants

The trial includes several companies, all members of the SVA as listed in Table 8. Content was kindly provided by the VR Industry Forum.

Role	Company	Details	Note
Content	VR-IF	F3 content 4K and 8K	Alternate 4K and 8K source
Encoding	Harmonic	Encoding and packaging of content to different formats for 4K and 8K viewport	
Networks	Orange 10G Spectrum 400M FiOS 1G	Commercial services	
CDN	Qwilt	CDN that can support open caching	Open caching not tested in the paper
Devices	Viaccess-Orca, Lumen Verizon	Quest 2 Galaxy S10	Alternate 4K and 8K source
Player	Viaccess-Orca	Embedded player in Viaccess-Orca Player Sample Unity App for Oculus Quest 2 and Android devices	Alternate 4K and 8K source
Analytics	Nice People At Work	Youbora test suite integrated in Viaccess-Orca Secure Player	

Table 8 - Participants in the SVA trial

Production Workflow

Figure 4 shows the workflow used to produce, encode and stream the content.

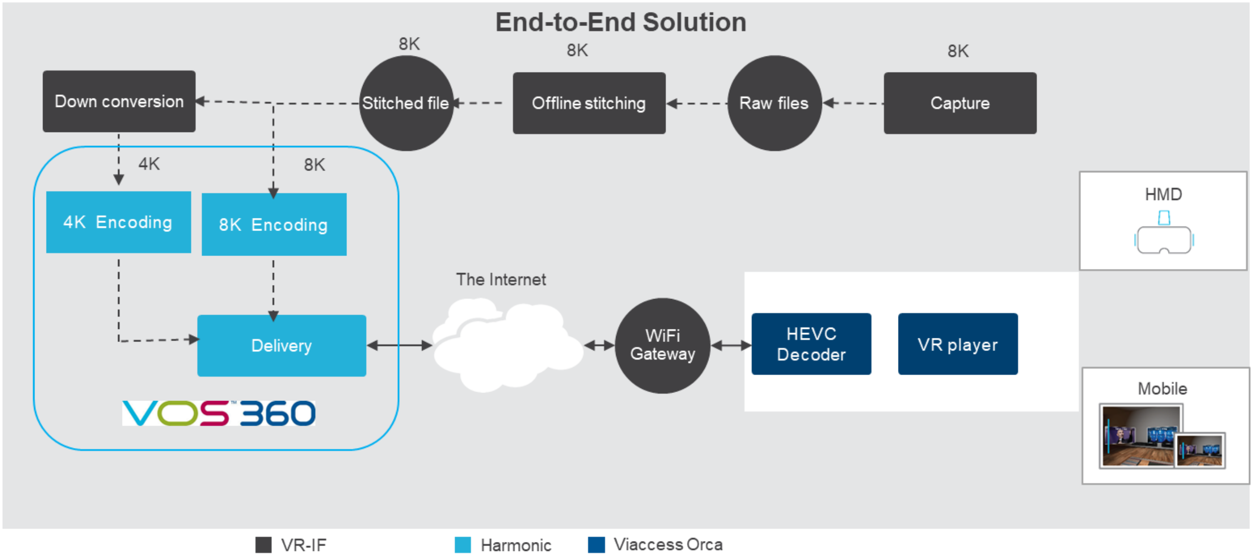

Figure 4 - Production workflow

The content is captured on an 8K multi-cam rig and is stitched offline. After stitching, we end up with a mezzanine file representation in 8K format (7680x4320x60). This file is then encoded in ABR mode at two different resolutions, 4K and 8K, using the Harmonic VOS360 cloud streaming platform. The reason we have two content representations is to measure the impact on HMDs when streaming 8K vs. 4K content. We also want to see the difference when 4K content is played on 8K devices as opposed to legacy 4K ones.

Once the content is encoded, it can be streamed over the internet and decoded by the Viaccess-Orca Secure Player.

Network Setup

The first phase of the trial was executed using a standard CDN provided by Qwilt combined with a very-high-speed network.

The very-high-speed networks tested were:

- Spectrum DOCSIS 3.0 with 400 Mbps download
- Orange fiber with 10 Gbps download
- Verizon FiOS 1Gbps download

Test Plan

The complete end-to-end workflow was used to test the different configurations listed in Table 9.

ID	Source	Device	Description
#1.2.1.1	4K	8K HMD	Test 8K HMD impact

#1.2.2.1	4K	8K phone	Test 8K phone impact
#1.3.1.1	8K	8H HMD	True 8K HMD experience
#1.3.2.1	8K	8K phone	True 8K phone experience

Table 9 – Tested configurations

The main purpose of the tests is to see first the overall performance of the system in a subjective way as well as in an objective way, comparing the different permutations of the content source (i.e., 4K vs. 8K).

Testing Methodology

The subjective testing was performed running all of the different tests in Table 8 while the objective testing was done using the Youbora analytics tool from Nice People At Work (NPAW). The NPAW objective score involves measuring the "happiness" score that is based on the calculation of: join time, buffer ratio, average bitrate played and buffer time.

The reported scorings range from 1 (worst) to 5 (best). We also report on the average bandwidth (of the served profiles) and the throughput (rate of request served to the client).

Testing Results

Table 10 provides the results of the testing done for the different scenarios, using the following networks:

- Spectrum DOCSIS 3.0 with 400 Mbps download
- Orange fiber with 10 Gbps download
- Verizon FiOS 1Gbps download

Note that the Wi-Fi connection to the device is limited to 400 Mbps.

ID	Network	Average bandwidth	Throughput	Subjective score	Objective score	Note
#1.2.1.1	Spectrum	15.9 Mbps	23.9 Mbps	3.0	3.8	
	Orange	15.9 Mbps	69.1 Mbps	3.5	3.86	
	FiOS	13.8Mbps	23.5 Mbps	3.5	6.4	
#1.2.2.1	Orange	15.9 Mbps	26.7 Mbps	4.25	3.84	
#1.3.1.1	Spectrum	26.1 Mbps	36.6 Mbps	4.5	4.5	

	Orange	27.7 Mbps	67.7 Mbps	4.5	4.62	
	FiOS	21.3 Mbps	25.9 Mbps	4.5	7.1	
#1.3.2.1	Orange	14.2 Mbps	26.5 Mbps	2.5	2.37	Player has lower performances vs. 4K

Table 10 - Test results

The table must be interpreted across two dimensions:

- the resolutions of the content
- the ISP used to run the tests

First, the performance of Orange, Spectrum and FiOS networks are very consistent for HMD testing, second, the result indicated a higher objective score on FiOS network. As the bandwidth measure vs other networks is the same, we can only explain the difference in objective score by the fact on the FiOS network, the highest bitrate is sustained more often.

Comparing HMD experiences (1.2.1.1 vs. 1.3.1.1), we see a difference of 35% on the subjective score and 15% on the objective score between 4K and 8K content. This is explained by the fact that the average bandwidth available is high enough for the 8K profile to kick in, and that users prefer the 8K experience.

Comparing mobile experiences (1.2.2.1 vs. 1.3.2.1), we see a good result on 4K which can be explained by the fact that on a mobile screen, 4K resolution is “good enough,” equivalent to a 2D 1080p experience. With 8K content we see a lower quality of experience (both subjective and objective) due to the fact the current version of the mobile player does not request more than 4K profiles.

Next phases

The next phase of the trial will be to measure the impact of open caching, deployed on Verizon FiOS network, and measure the impact on the QoE. This will be published in an upcoming SVA communication.

Conclusion

This paper described an 8K VR viewport-independent solution that can be deployed on very high-speed broadband networks using 8K-capable devices such as 5G phones and recent HMDs. Backward compatibility with legacy 4K devices is achieved by the virtue of ABR. The DASH environment offers continuity with existing OTT platforms that are already using the DASH protocol. The trial was using off-the-shelf OTT technologies such as NPAW analytics and DRM using CENC encryption. We have measured the performance of the solution on live fiber/DOCSIS networks and have found the technology fits quite well with the very high-speed network capability. We measured a significant improvement in QoE when moving to 8K vs. 4K and found that the 4K experience on mobile was already satisfactory. On the power dissipation side, with Quest 2 we see nearly no power consumption difference between 4K and 8K content, which makes the 8K VR viewport-independent solution green. The next step is to measure the impact of open caching on the QoE on FiOS fiber network. Next, we plan to test the technology with more features, such as DRM, low latency for live as well as on 5G networks, and see how the technology can unleash new fields of applications.

Acknowledgments

The authors thank Glenn Goldstein from Lumen and Sanjay Mishra and his Verizon colleagues for their contribution to the paper and for the testing of the solution within the Streaming Video Alliance VR trial.

References

1. R. Koenen, T. Fautier. 2017. Streaming UHD-quality VR at realistic bitrates: mission impossible? NAB 2017.
2. Virtual Reality Industry Forum (VRIF). June 2020. VRIF Guidelines 2.2. https://www.vr-if.org/wp-content/uploads/VRIF_Guidelines-2.2.pdf.
3. Harmonic. March 2016. EyeQ: achieving superior viewing experience. https://info.harmonicinc.com/technical-guide/achieving-superior-viewing-experience/
4. T. Fautier. November 2019. 8K is making progress bit by bit. Harmonic. https://www.harmonicinc.com/insights/blog/8k-making-progress.
5. Harmonic. December 2018. DASH CMAF LLC to play pivotal role in enabling low latency video streaming. https://info.harmonicinc.com/white-paper/dash-cmaf-role-in-enabling-low-latency-video-streaming/.
6. J. Chandaria, January 2021, A comparison of the carbon footprint of digital terrestrial television with video-on-demand: https://www.bbc.co.uk/rd/publications/whitepaper189

