CASE STUDY

Virtual Reality Streaming

Tiledmedia Tiled Streaming Enables Streaming 8K 360VR

IBC 360 Live system initiated by Tiledmedia and Intel live streamed the five-day IBC 2019 conference globally in 8K, 360 degree virtual reality (VR). Use of tiled streaming technology reduces bandwidth requirements making it a key enabler of high-resolution 8K VR streaming.

A new milestone for streaming virtual reality services was set at the IBC 2019 exhibition when the five-day conference, including the complete half-day Intel® Visual Cloud Conference, was streamed live to a global audience in 8K 360 degree virtual reality (360VR). To transmit in 8K 360VR in full would take a lot of bandwidth—between 60 Mbps and 100 Mbps of data to be decoded on the mobile device/headset. The project utilized tiled streaming to reduce both the bandwidth and decoding challenges.

Challenge

Minimizing the bandwidth needed for an 8K VR demonstration with no reduction in responsiveness, and giving access to 8K source content on devices with a 4K decoder.

Solution

Tiledmedia's ClearVR Cloud Live encodes the high-resolution video in tiles, where the client application with ClearVR integrated retrieves only those tiles that are in the user's actual viewport, with updates as they change their view.

Results

High-resolution 8K 360VR with very fast response to head movements and changes of view.

Achieving Good Quality VR Distribution

High quality, 8K VR images require a very large number of pixels, which impacts all aspects of producing and streaming an 8K 360VR video. 360VR is usually created using a number of video cameras that all record a piece of the environment. Stitching software then combines these video images into a single spherical video that is then projected onto a flat surface in the form of an equirectangular projection (ERP). A flat map of the earth is an example of what an ERP looks like. The ERP includes the entire image, but the VR user only sees about 1/8th of the complete picture at any one point in time through the viewport of their VR headset. Figure 1 shows the principle of how much of an entire image a viewer sees at any one time. The other parts of the image come into view as the user turns their head.

Since the video is played right in front of the user's eye, magnified by special lenses, the resolution needs to be very good. The industry standard in VR is to use 4K video for the entire ERP, which results in a 4,096 x 2,048 pixel image. But at any point in time, the user will only see an image through the viewport sized at approximately 1,000 pixels x 1,000 pixels per eye.

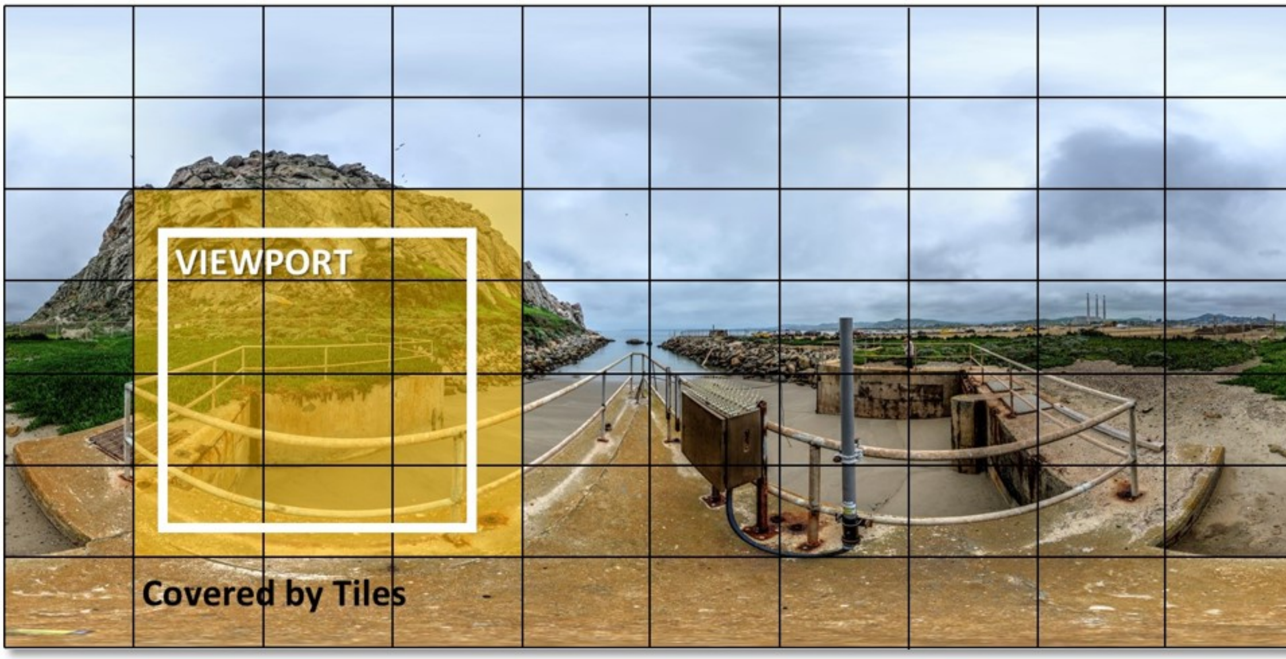

Figure 1. Image shows the viewport that is present on a VR device as a portion of the overall available image. Tiled streaming serves just the tiles needed for the viewport.[1]

This low resolution is increased when using 8K video. An 8K VR stream generates an 8,192 x 4,096 pixel image, yielding a viewport of about 2,000 x 2,000 pixels per eyes—considerably better.

The challenge is the amount of bandwidth required to get that high-quality image to the user, and the fact that existing devices cannot decode 8K video. Each image is made up of a huge number of pixels that are equivalent to 16 high definition TV channels, and videos of this resolution require network throughput of between 60 Mbps to 100 Mbps. In addition, this amount of data must be transmitted into the user's wireless VR device over Wi-Fi, and that device also has to have an 8K hardware decoder to decode the signal. Currently, there are very few devices that can decode 8K video.

Tiling the images, and then requesting only the image portion that is being viewed, is one way to reduce the bandwidth and make 360VR streaming services available on a broad basis. To demonstrate the potential of this approach, Intel and Tiledmedia and seven other partners set up the IBC 360 Live end-to-end chain, a pioneering live 8K VR distribution with worldwide availability. IBC 360 Live integrated hardware, software, and services that are all commercially available today. The other companies that contributed to the project included Akamai, Google, IBC, Iconic Engine, KPN, Oculus, and Voysys.

Tiledmedia ClearVR Cloud Live

The foundation of the distribution system is Tiledmedia's ClearVR Cloud Live, a content encoding and packaging service that is deployed on the Google Cloud and delivers low latency, high fidelity streaming services. ClearVR Cloud Live does this using tiled streaming.

The ClearVR Cloud Live service only requires choosing the quality and the output location, then contributing the stitched video to the cloud. The contribution streams are usually HEVC-encoded at very high, mezzanine bitrates.

Tiled Streaming Benefits

- Extremely high quality
- Virtually zero motion-to-photon latency (motion-to-photon latency depends on device; ClearVR does not add latency)
- Playback is available on either dedicated head-mounted VR devices or flat devices such as phones, tablets
- Uses standard encoding/decoding systems
- Can be used for on-demand and live content
- Offers global scalability to millions of users simultaneously over any CDN, using standard http streaming technology with no per-user edge processing required
- Bitrates for 8K video are reduced to be comparable to the usual VR distribution resolution of 4K

IBC 360 Live: Smart VR Delivery Through Tiles

At IBC 2019, a coalition of companies in the streaming VR value chain developed the IBC 360 Live distribution system to deliver a streaming consumer 8K 360VR service using tiled media to reduce image size and reduce necessary bandwidth and decoding compute power. The real-time nature of the experience, including the need to instantly serve up new content when a user's head moves, makes this approach particularly challenging. When the user turns their head, the system retrieves new high-resolution tiles, decodes, and displays them–all within a tenth of a second. This so-called "motion-to-high-resolution" latency obviously depends on network conditions. Under the wide range of conditions that we observed during the IBC 2019 project, over 63% of such tile switches happened within that tenth of a second.[2]

The production of the IBC 360 Live event was conceived and realized in less than two months, which was only possible because it utilized commercially available components. The encoded and distributed video was compliant with the high-efficiency video coding (HEVC) H.265 standard, which is implemented in hardware in all modern VR client devices. Tiledmedia is currently working in the Moving Picture Experts Group (MPEG) working group to contribute its tiled streaming innovations to the upcoming Omnidirectional MediA Format (OMAF) version 2 standard, planned for approval by mid-2020.

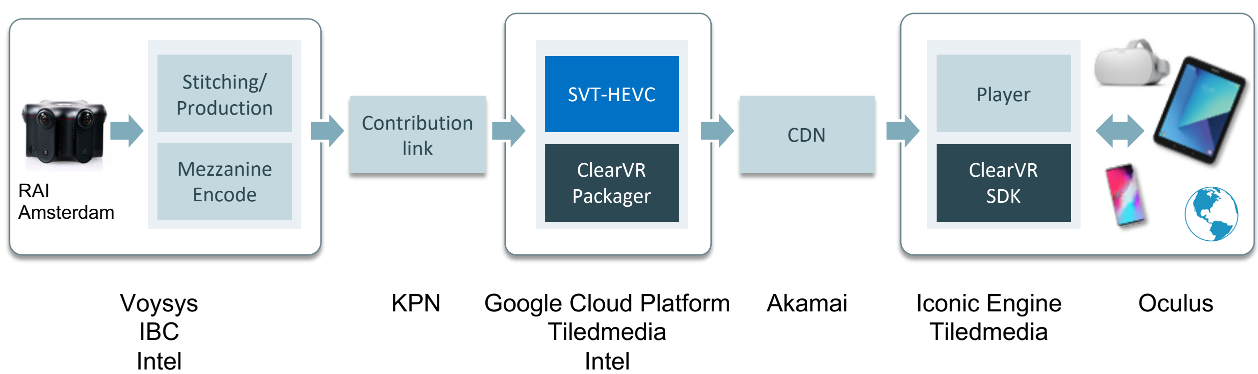

Figure 2. IBC 360 Live project overview

Tiled streaming was at the heart of the end-to-end IBC 360 Live demonstration, which was powered by the following components and processes:

- Two Kandao Obsidian R high-resolution 360VR cameras were used to digitally capture the event. Each of the VR cameras produced three individual feeds that were stitched in real-time into a single 360 degree sphere.
- The images were stitched using Voysys VR Producer software into a single equirectangular projection (ERP) with a resolution of 8,092 x 4,046 pixels. Voysys VR Producer was also used to mix in an image of the conference speaker slides as a virtual screen into the scene. The software converted the ERP into a "cubemap," where the image is mapped to the six faces of a cube. The is a format that is more suitable for tiled distribution. The cubemap has a resolution of 6,144 x 4,096 pixels, which preserves the 8K resolution along the equator of the virtual sphere.
- Two servers with Intel® Xeon® Scalable processors and advanced graphics cards were used to run the Voysys software and perform an initial HEVC mezzanine encode for each of the cubemaps. The mezzanine encoder lightly compresses the camera feeds for distribution to the cloud data center where the feeds are processed by ClearVR Cloud Live for tiled distribution.
- The two mezzanine-compressed ERPs were sent to the Google Cloud Platform (GCP) instance utilizing a dedicated 400 Mbps fiber link to the internet, supplied by KPN for the demonstration.
- Image cubemaps are input into the ClearVR Cloud Live software, which then decodes the six faces of each cubemap before dividing the decoded video into 16 tiles per cube face for a total of 96 tiles. These tiles are independently coded and stored on a content delivery network (CDN), where they are available for the client device.
- The HEVC-encoded tiles were then packaged by ClearVR Cloud Live into MP4 files for distribution. The processes to create the cubemap, and then tile, encode, and package the full 8K content run on multiple hundreds of Intel Xeon processor cores, in parallel processes managed by the ClearVR Orchestrator. ClearVR Orchestrator ensures all tiles are ready in time for ingest into the CDN and provides for redundancy and fail-over in the encoding and packaging. The total bitrate of the packaged files amounts to about 120 Mbps for each camera feed (numbers are doubled because of the two cameras used in the demonstration).
- Next, the packaged MP4 files were ingested into an Akamai CDN instance using six parallel Media Services Live (MSL) ingest points per camera, where each of these ingests can handle 45 Mbps. These parallel ingests were designed to accommodate the exceptionally high 120 Mbps bandwidth requirements, and allow for variation in the individual ingests. The multiple-ingest configuration was developed by Akamai and Tiledmedia in 2018 to accommodate live 6K and 8K VR events. As the IBC 360 Live production used two parallel 8K cameras, 12 (2 x 6) MSL ingests were required. The ingested files were then

recombined and uploaded onto Akamai's origin for its EU region and, from there, distributed across all regions globally.

- To distribute live or on-demand VR content, the Akamai network uses a VR-specific configuration that optimizes the availability of the requested tiles at the edge cache. This is important in providing the optimal response when users move their head or swipe on a flat device. This configuration relies on the http/2 or QUIC (http/3) protocol using with multipart byte range requests and smart pre-fetching of tiles to the edge cache. With tiled streaming, a short request/response delay is much more important than for regular on-demand services. In the IBC 360 Live demonstration, the "motion-to-high-resolution" latency was typically 3 frames or less at 30 frames per second. The ClearVR system logs performance statistics. Using these statistics, and totaling all sessions across the world, we could see that over 63% of high-resolution tiles were in the user's viewport within these 3 frames after head motion.[2] This performance is good enough to be perceived as virtually instant to the viewer.
- The delivery and decoding of the VR images was handled by a special application developed for the IBC 360 Live demonstration by Iconic Engine, an end-to-end provider of extended reality (XR) solutions. The application was developed for the Oculus platform and then extended for Apple and Google devices. This allowed playback of the VR streams on "flat" devices and in low-cost, cardboard viewers. Iconic Engine's application platform integrates Tiledmedia's ClearVR library, which enabled Iconic Engine to develop the IBC 360 Live app in a matter of weeks. In the IBC 360 Live app, the ClearVR technology understands what the user sees and then retrieves the required tiles for that viewport. ClearVR then reassembles the bitstream snippets for these individual tiles into one single legal bitstream, which is decoded by the HEVC decoder in the user device. The last step is putting all the tiles in the right place on the rendered sphere. All of this happens at the last possible moment, in real time, giving the user the highest possible quality on both headsets and flat screens. In the demonstration, the typical end-user bitrate varied between 12 Mbps and 17 Mbps with an average bitrate of 14.5 Mbps.[2] The end-to-end latency was measured to be about 30 seconds.

IBC 360 Live by the Numbers

Number of cameras: Two, user-switched

Event Duration: 5 days; 8+ hours per day

Capture resolution: 8,092 x 4,046

Contribution bitrate: 150 Mbps per camera (HEVC)

Distribution resolution: 8,192 x 4,096

Distribution Encoder: Scalable Video Technology for HEVC (SVT-HEVC)

Distribution Format: MP4-based ClearVR packaging

Distribution bitrate: 12 Mbps – 15 Mbps

Streaming Protocol: Standard http/2 with multipart byte range requests

User device decoder: HEVC Main level 5.1

Glass-to-glass latency: ~30 seconds

Supported devices: Oculus headsets, iOS and Android tablets and phones

Cloud processing: More than 1,000 Intel processor cores in Google Cloud Platform for the two streams (dynamically managed)[2]

Local Servers: MSI Trident X plus (MS B926)
Intel® Core™ i9 9900K (microcode 0x9E)
32 GB RAM (2x 16 GB DIMMs)
GeForce RTX 2080 TI
SSD Samsung NVMe M2 970 EVO plus (500 GB)
Seagate Barracuda 2 TB

Conclusion

Streaming of 8K 360VR provides the opportunity for new consumer and business entertainment and information services with realistic, high resolution images. Tiled streaming offers a way to dramatically reduce the bandwidth and decoding challenges involved to transport and consume the VR images. As shown in the IBC 360 Live demonstration during IBC 2019, tiled streaming delivers on the promise of highly responsive services that can be distributed globally.

For More Information

Visual Cloud Computing with Intel® technology:
www.intel.com/visualcloud

Tiledmedia ClearVR overview:
www.tiledmedia.com/index.php/clearvr-overview

[1] Figures provided courtesy of TiledMedia.

[2] Demonstration conducted by Tiledmedia and Intel. Performance measured in September 2019. See IBC 360 Live by the Numbers for configuration details. The following Google Cloud Platform machines were used for the project. For encoding: 3x n1-highcpu-96. For the GPU-accelerated projection conversion: 2x n1-standard-8 + V100 GPU. For the orchestration: 1x n1-standard-8. Configuration details for the GCP machines can be found at https://cloud.google.com/compute/docs/machine-types#n1_machine_type.