



AWS EC2 VT1 Instances Accelerate Video Transcoding with Xilinx

Amazon EC2 VT1 Instances Leverage Xilinx's Alveo™ U30 Media Accelerator Card to Deliver Best Price Performance for Multistream Video Transcoding with up to 4K UHD Resolution

AT A GLANCE:

Amazon Web Services (AWS) is the world's most comprehensive and broadly adopted cloud platform, offering more than 200 services from data centers around the world. The company serves a broad range of firms, from fast-growing startups to large enterprises and government agencies. AWS has the most-extensive global cloud infrastructure with 81 availability zones in 25 regions around the world.

Industry: Data Center
Location: Seattle, Washington
Established: 2006
Website: aws.amazon.com



SUMMARY:

As demand for video content increases, broadcasters are more frequently turning toward cloud services to reduce both transcoding and networking costs without compromising on quality or reliability. Amazon EC2 VT1 instances accelerate real-time transcoding and significantly reduce the cost of transcoding live video streams.

Powered by Alveo U30 media accelerator cards, AWS' EC2 VT1 instances offer the best price-performance for multistream video transcoding, with support for video streams up to 4K UHD resolution at 60 frames per second and the capability to transcode up to 64 1080p60 streams in real time per server. This powerful solution can further be scaled using Amazon Elastic Container Service (ECS) and Elastic Kubernetes Service (EKS).

CHALLENGE:

One of the most important stages of video stream distribution is live transcoding, where a live stream comes into a network and is prepared for distribution at mass scale. This transcoding process must be swift and maintain a high picture quality. AWS has a large array of customers running video transcoding workloads from TV networks creating live broadcasts to enterprise videoconferencing, e-sports, and distance learning.

Amazon’s EC2 service provides virtualized servers in the cloud. For transcoding, the company offers x86 servers, GPU servers, and Graviton Arm-based servers. Having already seen success with Xilinx for AWS F1, its general-purpose FPGA-as-a-Service offering, Amazon was looking to achieve even greater transcoding price performance with Xilinx’s Alveo U30 accelerators as well as offer more purpose-oriented instances for this important use case.

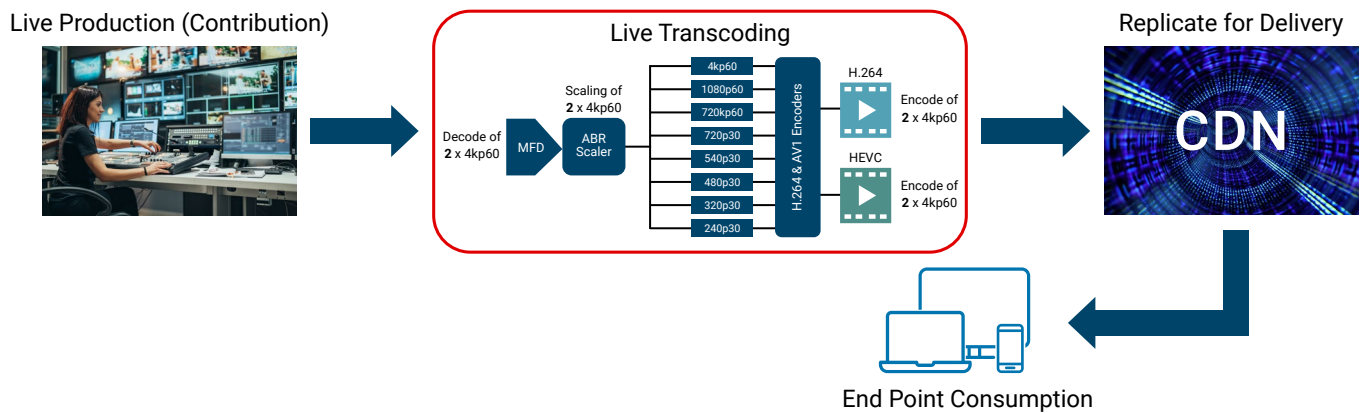


Figure 1 – Overview of a live streaming workflow

SOLUTION:

Amazon EC2 VT1 instances are built for video transcoding workloads, with up to eight Xilinx Alveo U30 media accelerators. The accelerated codecs in the Alveo cards support profiles that can be decoded by a broad range of end-user devices.

The Alveo U30 fits into a single PCIe slot and supports up to two channels of 4kp60 video streams. A single card can support up to eight 1080p60 streams and can be subdivided into lower resolutions down to 48 channels of 540p video resolution or below. The card supports H.264 and HEVC standards and operates at a very low 25 watts per card. It can decode and encode up to 48 video streams simultaneously.

The Alveo U30 platform was designed to support live transcoding, but it also supports adaptive bitrate scaling. Unlike other transcoding options, the U30 platform performs ABR scaling entirely on the card itself, without requiring CPU resources or cutting into the number of streams the card can support. This results in highly dense and efficient scaling. Another use case is faster than real-time (FTRT) transcoding. This leverages all the throughput of the card to transcode video streams at faster than real-time. For example, it can encode an 11-minute film in as little as 120 seconds, five times faster than real-time, just on one of the two devices on the card and is extensible across multiple devices and cards. With four devices (two cards), it could transcode the same film in around 38 seconds.

The Alveo U30 media accelerator supports all commonly used AVC profiles: Baseline, Main, and High as well as HEVC profiles: Main and Main Intra up to level 5.1.

RESULT:

Amazon EC2 VT1 instances, powered by Xilinx, deliver a lower cost-per-stream than the company's GPU- or CPU-based instances for live video encoding. They also support ultra-low video transcoding latencies for natural human interactions, and they can be bundled with AWS services to manage, scale, and package transcoding workloads.

"VT1 instances can provide up to 30% better price per stream than G4dn GPU-based instances, and up to 60% better price per stream than C5 CPU-based instances," said Leif Reinert, senior product manager at Amazon Web Services. "This is important because not only can customers save on infrastructure costs, but they can also enjoy the stream density and low latency they need without compromising on video compression or visual quality," he added.

The new instances are offered in three sizes ranging from one Alveo U30 card, 24 GB of memory and 3.125 Gbps networking speeds to eight Alveo U30 cards, 192 GB of memory, and networking speeds up to 25 Gbps., and they will be available on AWS Outposts soon.

ADDITIONAL RESOURCES: [Learn More About AWS's EC2 VT1 Instances](#)
[Learn More About Xilinx Alveo U30 Accelerator Cards](#)
[Learn More About Xilinx Video SDK](#)

Corporate Headquarters
Xilinx, Inc.
2100 Logic Drive
San Jose, CA 95124
USA
Tel: 408-559-7778
www.xilinx.com

Xilinx Europe
One Logic Drive
Citywest Business Campus
Saggart, County Dublin
Ireland
Tel: +353-1-464-0311
www.xilinx.com

Japan
Xilinx K.K.
Art Village Osaki Central Tower 4F
1-2-2 Osaki, Shinagawa-ku
Tokyo 141-0032 Japan
Tel: +81-3-6744-7777
japan.xilinx.com

Asia Pacific Pte. Ltd.
Xilinx, Asia Pacific
5 Changi Business Park
Singapore 486040
Tel: +65-6407-3000
www.xilinx.com

India
Meenakshi Tech Park
Block A, B, C, 8th & 13th floors,
Meenakshi Tech Park, Survey No. 39
Gachibowli(V), Seri Lingampally (M),
Hyderabad -500 084
Tel: +91-40-6721-4747
www.xilinx.com