AMAZON EC2 VT1 INSTANCE FOR LOW-LATENCY VIDEO STREAMING IN THE CLOUD

OVERVIEW

Global demand for video content has been rapidly increasing and now is the major driver of Internet and mobile network traffic. Over-the-top streaming services are seeing an explosion of content creators seeking to deliver great image quality, while live event broadcasters need to scale with demand while controlling infrastructure costs– all without sacrificing reliability.

One of the most compute-intensive stages of video distribution is transcoding—a critical step that optimizes video for different network bandwidths and device endpoints. The challenge of transcoding hundreds or thousands of streams at scale include accessing the necessary compute power (based on fluctuating demand), integration into existing production pipelines, and securely delivering video in real-time.

SOLUTION

To scale content delivery, broadcasters are more frequently turning toward cloud services to minimize networking costs without compromising quality or reliability. Amazon EC2 VT1 instances accelerate real-time transcoding while reducing the cost of delivering live video streams when compared to CPU- and GPU-based transcoding solutions.

Powered by AMD Alveo™ U30 media accelerator cards, the VT1 instance supports up to 16 4K UHD streams at 60 frames per second as well as 64 1080p60 streams. The instance can also be used with Amazon Elastic Container Service (ECS) Elastic Kubernetes Service (EKS), and other AWS services for orchestration, observability, infrastructure monitoring, scaling, packaging, distribution, and more. Customers save on infrastructure costs while maximizing low-latency stream density, all without compromising video quality and user experience.

HIGHLIGHTS

Lowest Cost per Stream on the AWS Cloud

- 60% lower cost/ stream vs. Amazon EC2 C5 CPU-based instances†
- 30% lower cost/ stream vs. Amazon EC2 G4dn GPU-based instances†

Low Latency Transcoding that Balances Video Quality and Network Bandwidth

- Deliver immersive watching experiences for interactive media at up to 4K resolution
- Adaptive bitrate scaling for optimal viewing experience while minimizing bandwidth cost

Turnkey Solution for Ease of Deployment

- Support for familiar frameworks (FFmpeg, Gstreamer) and a C-API for full customization
- End-to-end AWS services including management, scaling, packaging, distribution, and more

60% **Lower cost-per-stream vs. Amazon EC2 C5 CPU-based instances†**

KEY APPLICATIONS

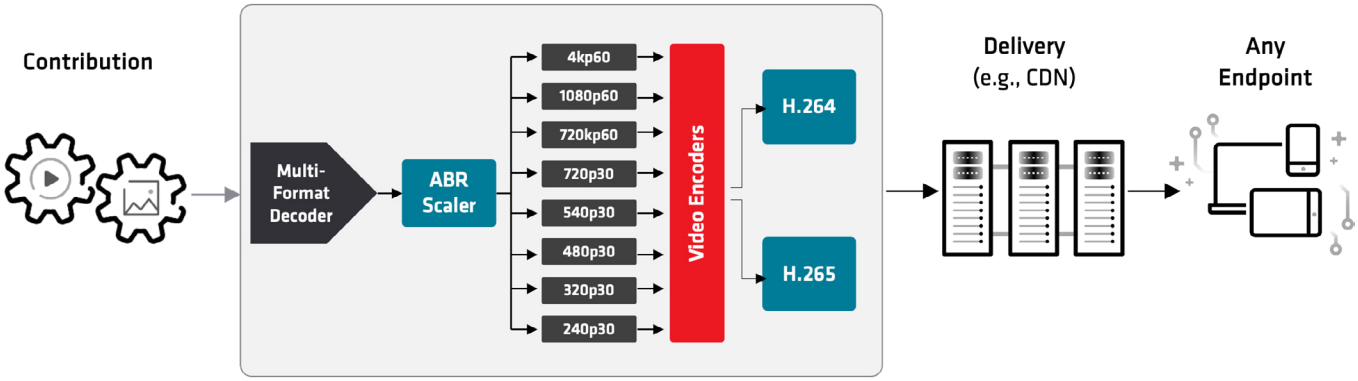
- Video conferencing
- Virtual Events
- Watch parties
- Telemedicine
- E-learning
- Live auction
- Live sports betting
- Smart homes
- Infrastructure security

LOW LATENCY TRANSCODING BALANCES VIDEO QUALITY AND BANDWIDTH

Each Amazon EC2 VT1 instances is powered by up to 8 AMD Xilinx® Alveo™ U30 media accelerator cards, with each card supporting up to 16 4k60p streams that can be subdivided into lower resolutions–down to 64 1080p60 streams. The instance delivers the following:

- **Low latency transcoding,** enabled by AMD hardware acceleration technology optimized for the H.264 and H.265 (HEVC) standards. Multiple transcoding jobs are managed across a card for deterministic latency, with multiple cards per instance for ease of scalability.
- **ABR** (adaptive bit-rate) **scaling** where a single input stream can be scaled down to a collection of lower resolutions and sent over a content delivery network (CDN) to ensure viewers on any endpoint can have continuity of experience regardless of network conditions.
- **Faster Than Real Time (FTRT)** transcoding for file-based use cases, where content providers can split raw file into multiple segments to encode in parallel, resulting in a fully encoded video file in a fraction of the time vs. traditional methods.

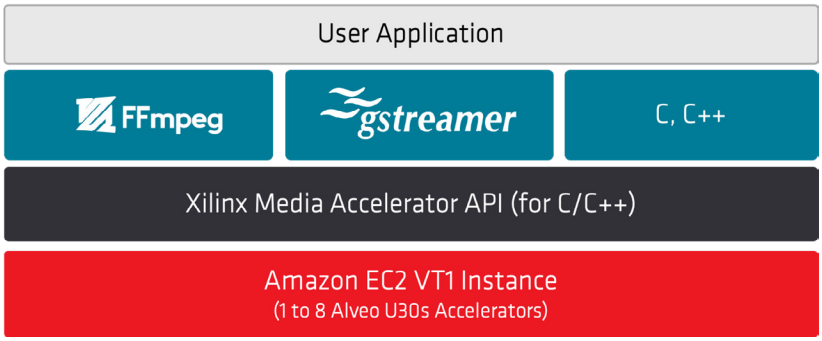
Hardware Accelerated Transcoding
on Amazon EC2 VT1 Instance



SUPPORT FOR FAMILIAR FRAMEWORKS AND C-API FOR FULL CUSTOMIZATION

VT1 instances integrate standard media frameworks for ease of migration of existing applications. The AMD Xilinx® Alveo™ U30 Video SDK is complete with tutorials, example designs, and a quick start guide–all available on GitHub. With FFmpeg and Gstreamer support, developers can get started quickly using pre-compiled libraries. Further customization is also available with a full-featured C API. Users can launch a VT1 instances using Xilinx Amazon Machine Images (AMI) on AWS Marketplace to rapidly test their own use case in the cloud.

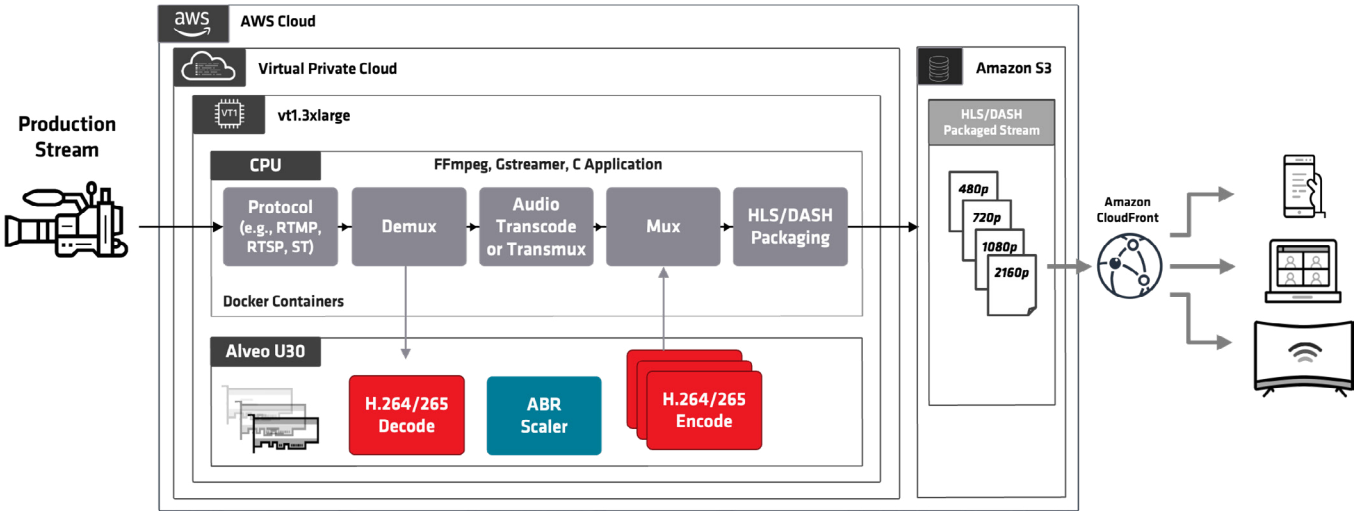
The Alveo U30 Software Solution Stack
Running on Amazon EC2 VT1 Instance



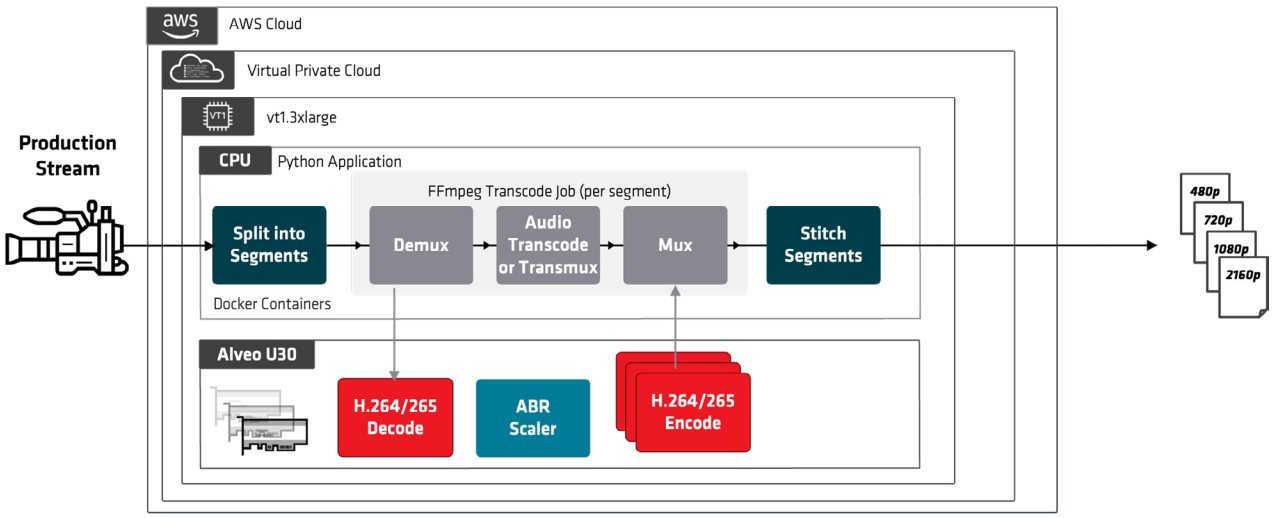
END-TO-END AWS SERVICES TO DEPLOY YOUR STREAMING CONTENT

Content providers can use VT1 instances with other AWS services to manage, scale, package, and prepare their transcoding workloads. Customers can manage and scale via Amazon Elastic Container Service (ECS) and Amazon Elastic Kubernetes Service (EKS), as well as deliver the final content globally using Amazon CloudFront. VT1 instances can also be complemented with AWS Elemental MediaConnect for secure live video transport. The diagrams below represent example AWS configurations for both live video and file-based use cases.

Live Video Streaming Configuration



Faster Than Real-Time Configuration (File-Based Transcoding)



CONFIGURATIONS AND AVAILABILITY

Amazon EC2 VT1 instances are available in three sizes, all of which can transcode multiple streams per instance. The vt1.3xlarge instance provides a single accelerator card and offers pixel processing bandwidths up to 2 4k60p (or 8 1080p60) streams in real-time. The vt1.24xlarge instances provide 8 accelerator cards providing the capability to transcode 16 4k60p (or 64 1080p60) streams in real-time. VOD assets can also be segmented and transcoded faster than real-time.

- **World-Wide:** availability in the AWS US East (N. Virginia), US West (Oregon), Europe (Ireland), and Asia Pacific (Tokyo) regions

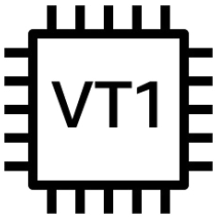
- **Instance Types:** where customers can purchase VT1 instances as On-Demand Instances, Reserved Instances, Spot Instances, or as part of Savings Plan.

- **Outposts:** VT1 instances will also be available on AWS Outpost racks to integrate their video feeds for transcoding at edge locations.

AMAZON EC2 VT1 INSTANCE PRODUCT DETAILS
(Full Specifications Here)

Instance Size	vCPUs	Alveo™ U30 Cards	Network Bandwidth	Memory(GiB)	1080p60	4Kp60 Streams
vt1.3xlarge	**12**	**1**	**3.125Gb/s**	**24GB**	**8**	**2**
vt1.6xlarge	**24**	**2**	**6.25Gb/s**	**48GB**	**16**	**4**
vt1.24xlarge	**96**	**8**	**25Gb/s**	**192GB**	**64**	**16**

NEXT STEPS



- Visit the Amazon EC2 VT1 instance page to view full specifications and pricing
- To learn more about the AMD live video streaming technology visit www.xilinx.com/livestreaming
- To inquire how the VT1 instance can work for your application, visit the product inquiry form

† https://aws.amazon.com/ec2/instance-types/vt1/
All performance and cost savings claims are provided by Amazon Web Services and have not been independently verified by AMD. Performance and cost benefits are impacted by a variety of variables. Results are specific to Amazon Web Services. GD-181

DISCLAIMERS

COPYRIGHT NOTICE