# LIVE VIDEO DELIVERY FROM THE CLOUD

A New Paradigm for Live Video Broadcasting

# CONTENTS

# INTRODUCTION

## A NEW WAY TO DELIVER LIVE VIDEO

Since the inception of traditional free-to-air television, there have been many technological and business-driven improvements in the media and entertainment industry to bring premium video content to viewers as efficiently as possible. These improvements include efforts as varied as the introduction of color television in 1953, the implementation of vast cable networks in the 1980s, the transition to high-definition (HD) broadcasting in the last decade, and the emergence of 4K resolution video in this one. Perhaps most important is the now-widespread implementation of advanced compression technologies and streaming protocols to enable video delivery to connected devices.

Consumers have embraced the explosion of over-the-top (OTT) video. Today, an estimated 1.6 billion people worldwide watch video on connected devices.[1] Seventy percent of them consume video on smartphones—twice as many as in 2012.[2] More than 60 percent of the video content people consume is watched live.[3] And, by 2020, half of all video will be consumed on a mobile screen.[4]

For content providers, the technology infrastructure required to create live streams for every consumer's device has been difficult to design and costly to procure and maintain. The advent of cloud-based live video processing heralds a new way to provide live broadcast programming to consumers. Now, cloud-based resources can support an end-to-end video processing and delivery solution for content distribution of live events and linear broadcasts.

## THE BENEFITS OF CLOUD-BASED INFRASTRUCTURE

In recent years, the cost of cloud resources has fallen as cloud infrastructure providers take advantage of economies of scale, increased network bandwidth, and the ability to offer enterprise-grade reliability and security. This offers content providers an opportunity to "get out of the infrastructure business," moving away from on-premises data center architectures to cloud-based resources for processing, packaging, and delivering live and on-demand video.

The transition to cloud-based video infrastructure promises several advantages for the production and distribution of live content, such as:

**Pay-as-you-go expense models**: Most cloud providers bill for their services on a pay-as-you-go basis. Customers pay no upfront capital expenses and need not invest in unneeded capacity or overprovisioned systems. For live events, producers can spin infrastructure up and down as needed and pay only for the resources they use.

**Scalability:** Cloud-based infrastructure can scale resources up or down automatically in line with demand, eliminating time spent provisioning resources and avoiding the need to guess capacity requirements or over-provision infrastructure for peak viewership. In fact, broadcast-grade video software

---

1 The OTT Playbook | Parks Associates

2 TV & Media 2017 | Ericcson ConsumerLab

3 TV & Media 2017 | Ericcson ConsumerLab

4 TV & Media 2017 | Ericcson ConsumerLab

operating in the cloud can readily address unexpected spikes in audiences for live broadcasts as demand surges.

**Operational agility:** The cloud allows low-cost experimentation and the capability to provision and deploy new video services within minutes. This supports rapid development of new live content or channels, allowing real-time adaptation to changing audience interests, and helps organizations quickly capitalize on new business opportunities. Connections to multiple content delivery networks (CDN) are easily made based on distribution needs. In addition, cloud-based platforms lend themselves to the use of modular services, such as machine learning (ML) tools, to quickly introduce new capabilities to the video workflow.

**Global reach:** Cloud providers can deploy resources like media storage, database services, video encoding, and media packaging, as well as CDN distribution, anywhere in their global footprint. This drives efficiency by allowing video to be ingested near the content owner and delivered close to the consumer, anywhere in the world. For media organizations that operate globally, this facilitates fast, cost-efficient delivery of live events and live linear channels into worldwide markets.

**Focus on core business:** By using cloud-based video services, organizations can set aside the undifferentiated heavy lifting of acquiring and maintaining video infrastructure, instead focusing their personnel and financial resources on creating and delivering high-quality content and pioneering video services that delight their viewers.

**Support new technologies:** As new device types, video codecs, and streaming formats reach the mainstream, cloud-based video infrastructure can stay ahead of accelerating technology trends without investments in software upgrades or costly hardware replacements. This extends to related capabilities, such as artificial intelligence and machine learning, that are beginning to integrate with video processing and delivery workflows.

## CONTENT SOURCING

Moving live linear or real-time content to the cloud has been one of the greatest challenges faced by video providers. Yet, this has also become significantly easier and more reliable over time. The process starts with real-time inputs: These typically originate with on-premises encoders, such as AWS Elemental Live, either single or paired for redundancy, which convert an input signal into a mezzanine format and deliver redundant outputs to the cloud for stream creation and OTT delivery. The latest technology advancements allow encoders to use high-efficiency video coding (HEVC) to deliver the signal over real-time transfer protocol (RTP) with forward error correction (FEC) to ensure the resiliency of the input signal to the cloud.

Today, content can be delivered from on-premises encoders to the cloud via dedicated IP network connections such as AWS Direct Connect, over the open internet, or through data center bridges between traditional fiber networks such as the Level3 Vyvx or Hibernia networks. Non-fiber cloud delivery methods have expanded from real-time messaging protocol (RTMP) to also include Apple HTTP Live Streaming (HLS) and RTP. Each of these methods has advantages and disadvantages; choosing which is best suited for a particular video workflow depends on a number of factors, including the type of content being supported, the video processing approaches used, and goals for reliability and latency. For example, RTMP has widespread adoption and is supported by many open-source encoding solutions. HLS has reliability advantages and can include embedded data from broadcast feeds by using segmented transport streams, but it increases end-to-end network latency and is HTTP-based, which is

prone to data transfer retry issues. RTP promises reduced latency, supports embedded data, and can enhance reliability with the use of forward error correction. And, new encoding compression options, such as HEVC, offer improvements in the quality of video content while reducing overall bitrate and bandwidth load.

The workflows supporting channels with file-based sources are unique. They need to be able to handle requirements directly within the encoder that would have traditionally been included in a playout server, such as the ability to remap audio and captions, overlay motion graphics, support ad insertion, and update the content playback schedule on demand.

## CONTENT PROCESSING

As today's video consumers demand more playback capabilities and augmented services, pay TV providers and content owners need video workflows that personalize the viewing experience at the click of a button. Demand is growing for both OTT as well as live-to-VOD content, in which video-on-demand (VOD) assets are generated directly from live streams, and consumers increasingly expect DVR-like functionality with streaming video, such as pause and rewind for live broadcasts. To stay competitive, content providers must offer a range of live-to-VOD capabilities including time-shifted TV, catch-up TV and network personal video recorder (nPVR) as a part of OTT service offerings. As machine learning tools enter the video workflow, live-to-VOD workflows can apply new capabilities that add significant value to live and VOD content. For example, production workflows are currently deployed to extract video metadata from live streams using Amazon Kinesis Video Streams and Amazon Rekognition.

Traditionally, these capabilities were only available to the largest media organizations with sizable economic and technical resources because it is extremely complex to deliver this functionality to consumers. The cloud, used in combination with broadcast-grade, software-based video solutions, such as AWS Media Services, a family of fully managed cloud video services, can enable media organizations of all sizes to meet customer expectations efficiently and cost-effectively for 24/7 linear content delivery with nPVR functionality.

Linear services with live-to-VOD capabilities must address a number of technical requirements. Content providers must be able to repackage content, using catch-up TV or nPVR content as deliverable mezzanine files, for the wide variety of devices that need to be supported. The implementation needs to be scalable, flexible, and highly available from ingest to delivery to the consumer. With premium content, it is also important that any video services incorporate digital rights management (DRM) technology to protect valuable content regardless of the device used for playback and support service agreements with content owners.

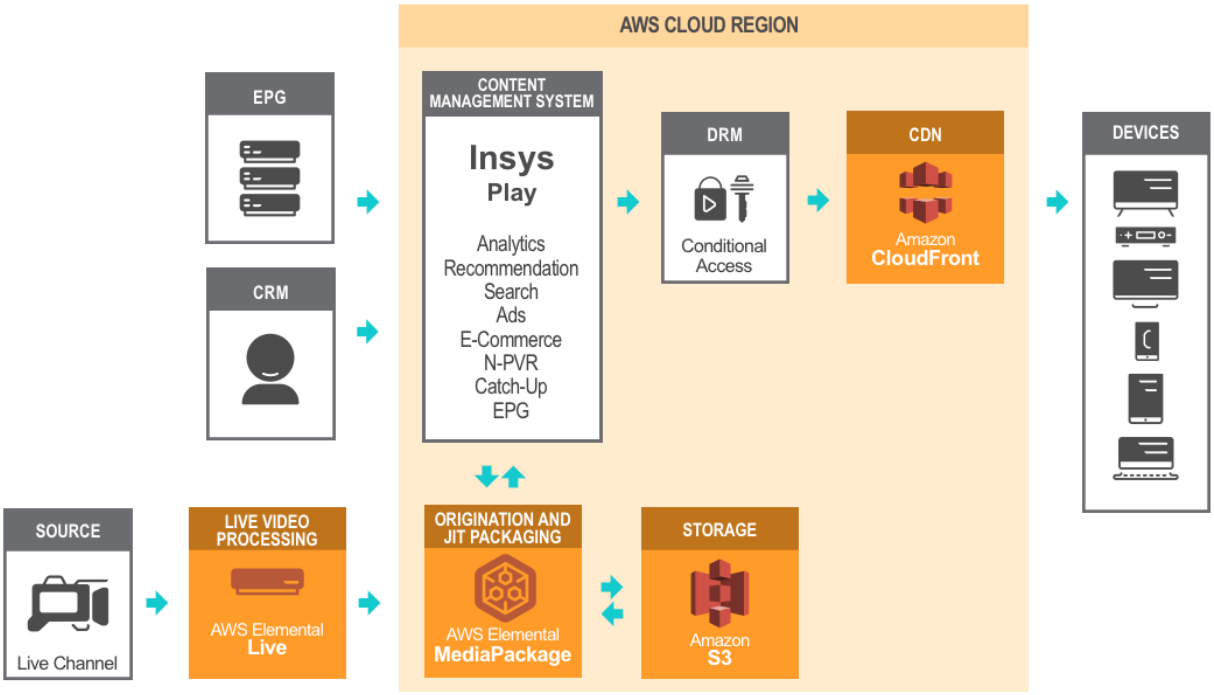


***Figure 1 – A video workflow supporting comprehensive live-to-VOD services***

Content providers need to ensure they are ready to adapt their services to new and updated viewing devices and streaming protocols. By relying on a software-based approach, broadcast solutions can be more easily upgraded to embrace new standards and features as they emerge. By also including a cloud-based, just-in-time (JIT) packager, such as AWS Elemental MediaPackage, that can adapt video streams to network and device parameters in real time, video providers can be prepared for whatever comes next.

## CONTENT DELIVERY AND MONETIZATION

Delivery of premium video to the user can be simple or complicated depending on content usage, user rights, and the intended monetization strategy. These considerations must be balanced with the best possible viewing experience.

After an adaptive bitrate (ABR) bouquet is created and video streams are ingested and stored, a cloud-based packaging service, such as AWS Elemental MediaPackage, can play a key role in content delivery. It is able to repackage content to match client requests for both live and live-to-VOD OTT delivery. Such services can detect the user's video playback device as well as network conditions and provide the appropriate encryption and streaming format to deliver optimized video streams. In addition, AWS Elemental MediaPackage works with a number of CDNs, including Amazon CloudFront, allowing video providers to deploy a multi-CDN strategy to help ensure a user receives content with the highest available service quality. And finally, these cloud-based packagers allow for alternative content options when geography-based content blackouts are needed.

With respect to monetization, cloud-based video workflows provide a range of options for transactional models (such as pay-per-view events), subscription models, such as Netflix, and ad-supported models,

such as fuboTV. As an example, cloud services with integrated content protection features, such as DRM support, can allow a content provider to offer their customers different levels of subscription access in alignment with their preferred monetization model. AWS Elemental MediaTailor, another of the AWS Media Services, performs content personalization and monetization and can work with AWS Elemental MediaPackage or another origin service. With server-side ad insertion approaches such as those enabled by AWS Elemental MediaTailor, video providers can insert personalized advertising into live video streams prior to packaging. The service calls upon ad decision servers to furnish an ad based on information about the user supplied by the client device and in accordance with the provider's preferences. Ads are presented to viewers as part of the primary content stream, with the same format and quality, which can help mitigate the effects of ad blocking software. Server-side ad insertion provides all the personalization capabilities without the quality or buffering issues created by client-side ad insertion, such as volume or format mismatches from content to ads. Audiences get a broadcast-like experience, while providers and advertisers are able to better engage viewers with more interesting and relevant advertising.

With machine learning tools increasingly integrated into video workflows, video providers can further augment their targeted advertising strategies by deploying automated ML services, such as Amazon SageMaker, Rekognition, Transcribe, and Comprehend, to build ad recommendation engines that may account for the viewing habits, engagement with previous advertising, content preferences, the current content being played, audio tracks, or other contextual factors to personalize advertising content.

## HIGH AVAILABILITY AND RESILIENCY

High availability is a critical requirement for any premium content. Traditionally, video providers have built redundant data centers in diverse geographic regions as a means of protecting assets and ultimately revenue streams. With the cloud, operators can now realize redundancy and high availability without the costs of duplicate hardware installations. For example, AWS infrastructure is built around Regions and Availability Zones. AWS Regions provide multiple, physically separated and isolated Availability Zones that are connected with low latency, high throughput, and highly redundant networking.

AWS Media Services take advantage of this framework with built-in high availability features. Resources for a live channel are spread across Availability Zones within a region, without users having to manage or monitor individual instances. In the event of a service failure, this dynamic, multilayered approach to availability ensures uninterrupted video processing and delivery. The reliability of services is designed to provide robust failover capabilities, with physically distributed resilient resources across multiple geographic regions. This model is true for the Amazon CloudFront CDN as well, providing a resilient workflow from ingest, into the cloud, all the way to the end user. Where video providers once had to procure and host additional servers for live video – if not a complete duplicate workflow – as failover targets, cloud resources can automate the process of maintaining high availability, without increased investment.

## USE CASE: SPUUL

Singapore-based Spuul is a leading VOD service provider, bringing more than 10,000 film titles to a worldwide audience across a full complement of TV and mobile platforms on a subscription basis. When Spuul launched a live streaming service, including 150 linear TV channels for viewers in Australia and New Zealand and a planned global expansion, it deployed an end-to-end live video workflow in the AWS

Cloud. The company's live offering incorporates a host of advanced OTT TV features, including catch-up and start-over functionality.

In the Spuul architecture, premium content is ingested from live feeds to the AWS Cloud via AWS Direct Connect and transferred to AWS Elemental MediaLive for real-time encoding. AWS Elemental MediaLive encodes adaptive bitrate bouquets, which are furnished to AWS Elemental MediaPackage for packaging in HLS and MPEG-DASH formats for ubiquitous playback across device types. AWS Elemental MediaPackage also applies encryption to the streams, addressing content security requirements. Packaged content is then delivered to Amazon Simple Storage Service (S3) to be stored for on-demand catch-up viewing, and to Amazon CloudFront content delivery network for low-latency distribution to subscribers' devices.

This use case brings to life much of the material touched upon earlier including premium content transferred as a single live stream and delivered to cloud-based encoders; real-time encoding of ABR bouquets, which are accessed by an origin server and packager for delivery of content to the end user on request; and built-in resiliency for seamless service. Several AWS services, including AWS Lambda serverless compute platform and Amazon Route 53 cloud domain name system, are also used in addition to those described above.

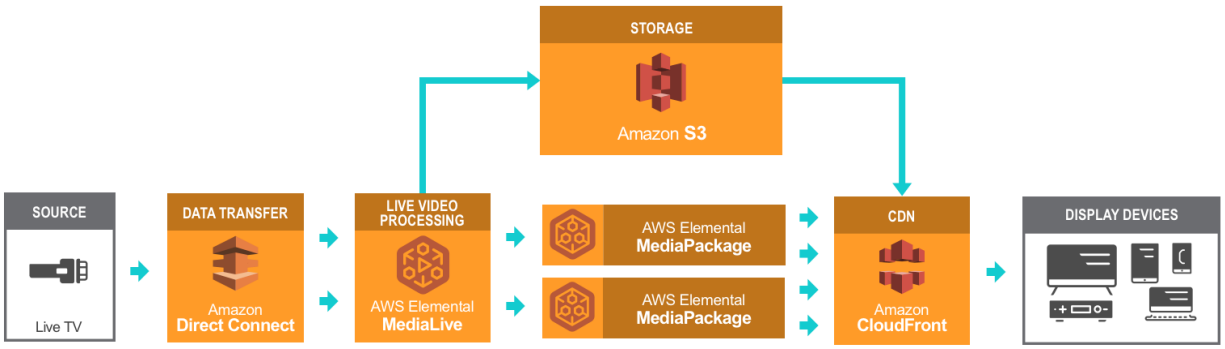


***Figure 2 – The Spuul video workflow for video-on-demand services***

## CONCLUSION: THE FUTURE OF LIVE TELEVISION

The broadcast landscape is changing. With end-to-end cloud deployments, content distributors can now outsource broadcast and video streaming infrastructure, with greater security, improved resiliency, and faster time to market.

AWS offers content distributors multiple benefits in support of delivering linear broadcast and live event content: rapid deployment of services to create multiscreen offerings with the same high-quality video processing capabilities as on-premises equipment; dynamic scaling of resources as demand fluctuates to accommodate spikes in viewership or number of live channels; secure protection of valuable media by creating resilient workflows with virtual resources that only incur costs when used; and the ability to enhance video content to provide content distributors with monetization opportunities through value-added features including time-shifted TV, dynamic ad insertion, targeted advertising, and DRM.

The migration of video workloads to the cloud is just beginning. Cloud video services from AWS have the power to fundamentally change how media and entertainment companies approach content processing and delivery. By moving operational costs to variable costs and paying only for what they use, content distributors now have the opportunity to experiment with enriched video offerings and pursue opportunities that resonate most with viewers. The conversation is no longer about how many systems need to be secured, racked and maintained to support an offering, but rather about how rapidly a new service can roll out, how many channels can be run, or how many VOD-asset hours can be offered per month to improve customer satisfaction. With the flexibility, scalability, and resiliency of the cloud, content distributors can offer new services to end users quickly, and implement and experiment with new features easily on a global scale, reaching more customers than ever before.

Stay up to date on the latest AWS Media & Entertainment news at:
aws.amazon.com/blogs/media