Networked Media. Working.

AMWA User Requirements for High Value, Low Latency, Live Video Production on Public Cloud

Introduction

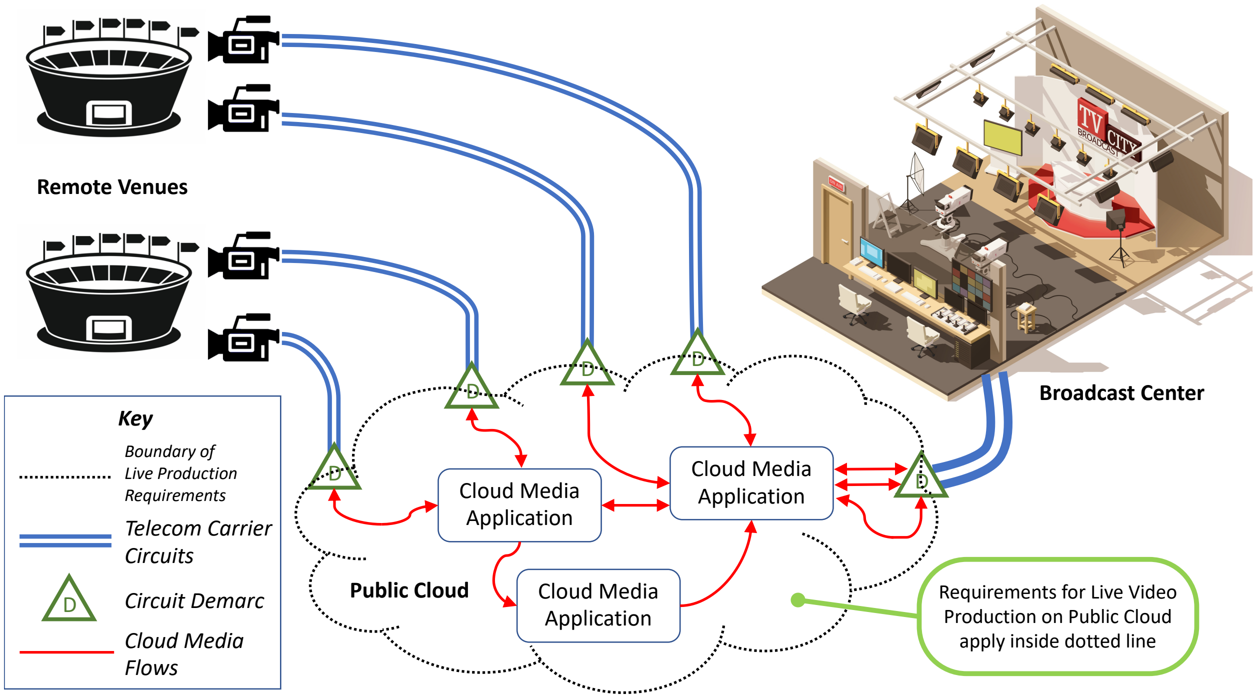
The video broadcast industry is gradually moving all aspects of its production and delivery processes away from bespoke electronic transports like SDI, and towards networked IP and scalable computing infrastructures. It has been demonstrated that the processing requirements of broadcasting infrastructure can be performed on COTS computer systems, and the industry desires to implement these processing functions in combinations of private and public cloud. The Advanced Media Workflow Association (AMWA) presents the following requirements for live video production on public cloud.

Scope

These requirements apply to high value live video production, intended for large audience live viewing. This includes major sports league matches and high-profile live news or entertainment, and could be viewed by hundreds of thousands to hundreds of millions of people. Low latencies in the media processing chain are essential for effective production of this kind of content. (There are other types of live cloud-based production that may have less stringent or different requirements.)

Business Requirements

High value live productions have a business requirement to deliver all content (especially paid commercials) flawlessly and with a minimal, managed amount of latency. Frames should never be added or dropped. This general requirement of managing latency and loss needs to extend from the telecom carrier circuit bringing broadcast media flows into the cloud, to and between the cloud-based applications hosted throughout the cloud provider's ecosystem, and back to the telecom carrier circuit carrying cloud-processed media flows to the broadcaster's premises. These business requirements drive the technical requirements in this document that apply inside the dotted line in the region of requirements diagram.



Region of Requirements Diagram

Communication Performance Guarantees

In on-premise systems today, broadcasting companies achieve near-perfect QoS through networks built to be inherently non-blocking of all traffic, or built to be non-blocking of expected media flows through the use of software defined networking (SDN) techniques. In order to achieve similar performance in cloud environments, communication performance guarantees must exist inside the public cloud, and also between the telecom carrier circuits and the applications running inside the cloud.

Media Transfer from the Ground to the Cloud

Broadcasters have a long history of partnership with telecom carriers in delivering high quality video with low latency and quasi-error-free performance. In support of the requirements above, a quasi-error-free path needs to exist from the telecom carrier circuit to the part of the cloud-based application which receives the incoming media flow.

Object Transfer Performance Guarantees Within the Cloud

In order to meet the latency and loss requirements above, cloud-based media applications require a mechanism which reliably transports media objects (for example, partial frames of video) between applications within a cloud provider’s ecosystem. The object transport needs to succeed within a (small) bounded period of time, in a quasi-error-free manner. Since realizable time bounding and the object size may be related, in order to minimize latency some applications may (for example) divide video frames into very small fragments – perhaps only a few lines of video – and expect transfer times well under 1ms. Other applications may choose to transfer objects as large as an entire frame, but will expect transfer times faster than an inter-frame period. For example, 16.6ms is the inter-frame period for 59.94 fps video.

Expected Media Flow Data Rates

High value live productions are likely to use "visually transparent" low-latency mezzanine video compression that maintains quality despite multiple coding/decoding cycles. Or such productions may simply use uncompressed video flows. For HD video flows at typical resolutions and frame rates, expected video data rates could range from 100 Mbit/s to 2.5 Gbit/s. Typical UHDTV1 (also known as "4K" or "UHD") rates could range from 1 Gbit/s to 10 Gbit/s. Typical UHDTV2 (also known as "8K") rates could range from 4 Gbit/s to 40 Gbit/s. Audio data flows will likely be uncompressed, at data rates of around 1.2 Mbit/s per audio channel. Additional flows of ancillary data, such as Closed Captioning, will generally have much lower data rates.

Media Transfer from the Cloud to the Ground

In support of the requirements above, a quasi-error-free path needs to exist from the cloud-based application which generates the final media flow to the telecom carrier circuit that brings the media flow out of the cloud to a broadcaster's premises.

Accurate Time Source

Broadcasters require an accurate time source within the public cloud to ensure synchronization of media flows. Broadcasters today use a combination of analog sync signals and IEEE 1588 Precision Time Protocol (PTP) within on-premise media networks and routinely achieve synchronization performance within 10 microseconds of International Atomic Time (TAI). The emission of the file replay, graphics generation, etc. needs to be frequency and phase aligned with other media flows in cloud-based applications in order to support the low-latency and zero frame loss objectives. This requires an accurate time source and the ability to wake the application at specific real times (for ad insertion, avoiding up- or down- cuts of next program, etc.) Time synchronization of cloud-based applications needs to be accurate to within 100 microseconds of TAI.

Efficient Point-to-Multipoint Distribution

Many media flows need to target multiple destinations. Within on-premise environments, multicast IP provided by Ethernet switches is a common solution to this requirement. In the case of cloud-based applications, the media flow arriving at the cloud from a telecom carrier circuit may need to be efficiently transported to multiple applications (for redundancy purposes, monitoring, recording, or simply as part of the production design). Cloud-based media applications may need to send media flows to several other cloud-based media applications, and to one or more telecom carrier circuits to exit the cloud. The overall efficiency of the infrastructure within the cloud service provider needs to accommodate this requirement.

Application Integration Requirements

The following sections describe requirements to allow for the rapid integration of existing media processing applications into live video production on public cloud.

Server-Based APIs for Rapid Integration of Legacy Software

Media infrastructure applications have been transitioning to software-based and COTS server processing environments for many years, and often make use of abstractions for media flows available within the common operating systems (such as DirectShow and similar abstractions). In order to facilitate the rapid adoption of cloud-based services, it may be helpful to provide APIs that are compatible with the media pipeline features that exist within the typical operating environments today. However, it is likely that in the future broadcast industry vendors can adopt more cloud-native solutions for their media processing applications.

Interoperability with ST 2110 Media Standards

Premise-based media infrastructures are increasingly built as SMPTE ST 2110 IP-based infrastructures. Appropriate engineering of the path from the ST 2110 infrastructure on-premise, through whatever type of IP transport mechanism gets it to and from the cloud, will be an essential part of the solution.

For further information or to ask questions, please use the Contact page on AMWA.tv
https://www.amwa.tv/contact

18-Apr-2020