Editing in the Cloud

By Ulrich Ening and Karsten Schragmann

Abstract

Changes in the way we work and globalization of the media industry have led to a growing interest in remote work options and collaboration—an area where the "cloud" offers promising opportunities. Cloud-based editing, for instance, presents a number of specific challenges such as how to provide editors with an on-premise (on-prem) feel and full client experience. In this article, we will look at how streaming technology and formats can be implemented together with an Adobe Premiere Pro client to enable users to work from anywhere without detriment to the user experience. Based on tests carried out together with end users, this article considers some of the technical innovations required to maintain that user experience, based on the use of streaming server and SMPTE RDD 25-based HD Proxies as a general approach to optimize the "work from anywhere" editing possibilities for two use cases addressing producers, journalists, and editors working remotely. We focus on (1) Editing with streaming proxy in a cloud-based Video Production Management Suite (2) On-premise installation aiming for a centralized system with different editing sites and remote access (3) A hybrid concept that combines the benefits of cloud and on-prem solutions.

Keywords

Cloud, content management, media asset management (MAM), media management, next generation, proxy- and craft editing

Introduction

An editing project is developed in the U.K., outsourced to Spanish animators, and then returned to the U.K. for further editing, such as adding audio and requesting additional work from the production team. The final version is delivered to a German customer who demands further changes; thus, the project needs to be edited in parallel in several countries to meet the date for delivery. This is just one example of how today's media industry relies on teams that collaborate across multiple locations.

Remote editing is complex and depends heavily on internet bandwidth. Editing projects consist of hundreds of source files, which can prevent fast delivery because content needs to be distributed across different sites. Recent broadcast surveys show an increasing trend in the importance of remote production (including remote editing). Despite the challenges, it has been a highly ranked trend in the broadcast and media industry in recent years.

For broadcasters, production houses—content creators in general—multisite workflows that include remote editing are challenging. Remote editing is complex and depends heavily on internet bandwidth. Editing projects consist of hundreds of source files, which can prevent fast delivery because content needs to be distributed across different sites. Recent broadcast surveys show an increasing trend in the importance of remote production (including remote editing). Despite the challenges, it has been a highly ranked trend in the broadcast and media industry in recent years.[1]

Traditionally, there have been two approaches to remote editing. First, "proxy" or "low-res" editing in which editors use specialized editing clients that utilize lower resolution and therefore less bandwidth. After editing, projects are either sent to "craft" editing clients, such as Adobe Premiere Pro, linking back to the high-resolution (HiRes) material, or a new clip is created—usually by a server-side render engine based on the HiRes material. This approach has many merits, especially when used for journalistic or highlight editing that requires only simple edits and/or voice over. But it is less suitable for other workflows as it offers fairly limited editing functions like lower quality that limits the evaluation of sharpness or depth of field. Editing clients specialized on proxy editing also do not offer the look, feel, and functions of craft editing clients.

A second approach, especially as high-bandwidth connections have become more widely available, has been to connect directly to the high-res storage with

Digital Object Identifier 10.5594/JMI.2020.2964181
Date of publication: 6 March 2020

a craft editing client over a virtual private network (VPN). However, even with high-speed broadband connections, this approach does not give users a local-like experience since the used protocols such as Simple Management Protocol (SMB), Apple Filing Protocol (AFP), and Network File System (NFS) are not really designed to operate via high-ping networks, resulting in high potential for packet loss, which dramatically degrades their performance.[2] Even audio elements take valuable time for the client to analyze and generate wave forms, often resulting in users disabling useful functionality just to make editing practical.

The following document examines an approach that is called cloud-based editing. Cloud technology, as defined by information technology (IT) analysts *Gartner* as "scalable and elastic IT-enabled capabilities delivered as a service using internet technologies"[3] can work as a foundation for collaborative and remote editing. But will cloud technology itself create an editing experience that comes close to established craft editing based on local content?

This article will discuss a cloud or hybrid solution that is based on the use of Adobe Premiere Pro combined with an optimized streaming server and SMPTE RDD 25[4] derived HD Proxies. HD Proxies in this case are high-definition H.264 compressed proxies, tailored for limited loss of quality. These proxies consider the technical innovations required to make cloud-based editing possible.

Cloud "Craft" Editing: What It Is—What It Is Not

Video editing is the process of editing segments of motion video production footage, special effects, and sound recordings in the post-production process.[5] This article addresses traditional assembly and video editing, also referred to as *craft editing*. This covers the editing done by producers and video editors. This type of editing requires a full video editing client like Adobe Premiere Pro or others. Other processes such as review and approval or client approval are not addressed in this article.

Challenges

Video editing presents a number of challenges that need to be overcome in order to create a full editing experience for editors or producers. On-premise software is installed and runs on computers on the premises—in the building—of the person or organization using the software. Those solutions provide fast encoding and playback of videos that are stored on local drives, storage or specialized production servers. Full quality high-res files are available to determine and evaluate the quality or sharpness. Fast forward/rewind as well as multicamera production or smooth playback with effects that are directly rendered into the playback is possible.

High Data Rates

HD and UHD video formats result in very high data rates, which need to be transferred via the network. Date rates

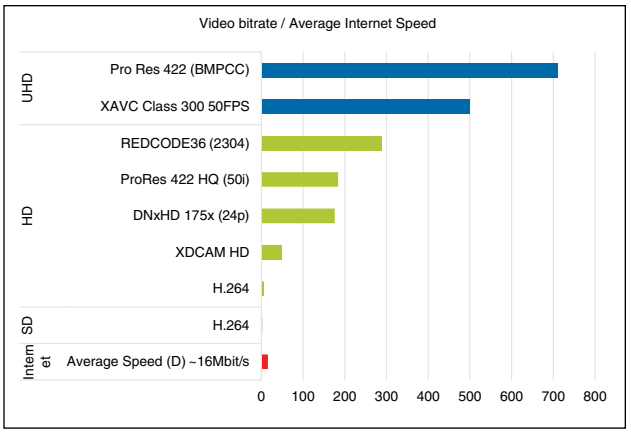


FIGURE 1. Video bitrate versus average internet speed.

of common formats range between 50 and 800 Mb/s, which already create challenges for on-premise installations typically consisting of 10 Gb network infrastructure and storage tailored to video production.

This comparison (**Fig. 1**) shows the average speed of the internet (in Germany)[6] compared with different production formats. As long as the average internet speed will not increase drastically, an editing workflow based on a format with less bitrate is necessary, which requires the use of a proxy file (proxy video as a low-bandwidth representation of a source clip).

Latency and Perceived Response

The bigger the distance between the client and the server, the bigger the latency of the data that needs to be transferred. This influences the perceived response of the editing client. A latency that is longer than the time of the displayed frame results in poor perception by the user. Depending on the frame rate, the maximum acceptable latency can be up to 15–20 ms.

Tools

Depending on the size and structure of a production company, the number of different editors working on different platforms or operating systems varies. Moreover, variations of hardware and software tend to grow. Thus, a solution is required that is able to flexibly deal with different variants. This could be achieved by using established and powerful craft editing software.

Security

How can media be protected from unwanted access? In an on-premise production environment, this can be achieved by blocking external access to the network. But the very nature of the cloud is to create maximum availability. As a result, a cloud editing solution requires protection via encryption or other security measures.

Costs

Costs of technology, necessary for an on-premise solution, need to be compared with that of a remote cloud solution. This results in a complex comparison, influenced by a variety of different cost drivers such as scaling,

utilization, and requirements. There is a choice between investment in hardware and continuous maintenance and support [capital expenditure (CAPEX)] or subscription of a software as a service (SaaS) cloud model with "pay as you use" subscription [operating expenses (OPEX)]. This article does not address the general on-prem versus cloud comparison, but suggests a technical streaming-based solution with a video bandwidth that is used by other distribution platforms like Netflix that streams 125 million hours of content per day.[7]

Methods for Remote Access

The following section compares three different approaches to achieving remote editing. They are based on different technologies that exist as remote editing solutions on the market.

Remote Display Control

A desktop environment (**Fig. 2**) can be hosted on a central system, including all available applications. This environment can be run remotely on another client. The client does not necessarily require full performance to run the hosted applications, but needs to be able to display the content delivered. Historically, the primary use of remote desktop software was remote administration. However, with the advent of cloud computing, the ability to provide sophisticated applications such as graphical editing, gaming, and video editing became more important. An editor accesses the system via a thin client based on protocols such as Remote Desktop Protocol (RDP), Virtual Network Computing (VNC), and proprietary protocols like PCoverIP. Proprietary protocols provide optimized delivery of content, like graphics or video, and include encryption.

An editing environment could be hosted on premise at a production house—including servers, storages, and editing client—or it can be completely hosted in the cloud. The range of editing possibilities can extend from simple editing to a large enterprise scale.

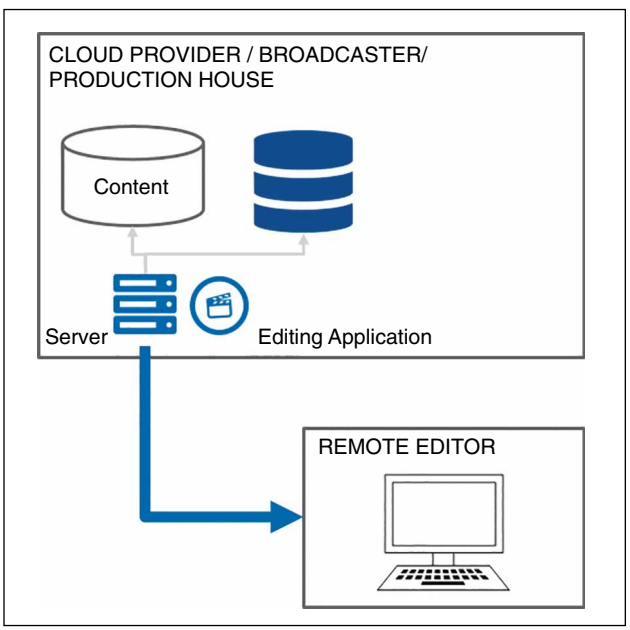


FIGURE 2. Remote display control.

Advantages

- The full feature set of an editing solution can be made accessible to remote users.
- Proprietary solutions are optimized for transferring video content smoothly and provide encryption.
- The editing solution can range from small, simple systems to an enterprise scale.
- Proprietary solutions allow displaying the application within a web browser.

Disadvantages

- The RDP and VNC protocols are designed for remote system access and administration. They will additionally compress the streamed content and will not provide higher frames per second. Thus, proprietary protocols are required to meet the requirements of perceived response.
- Latency is vital. Distance to the central system, hosted on premise, might be too great.
- Each editing instance requires a couple of centrally hosted servers and connected clients. The connection is 1:1. Scaling up the number of available editing clients requires the start of additional server instances.
- Scaling up additional instances might be limited to the licensing concept of the application vendor.
- Ingesting files cannot be performed via the editing client directly. The displayed application is passive. Other ways of adding material to collaborative editing are required.
- Limited depiction of content owing to compression and limitation of the frame rate.
- Solutions require high internet bandwidth to display video content smoothly.

Remote Editing—VPN

This remote editing approach (**Fig. 3**) utilizes functionalities that are already proven for editing solutions. An editing application runs on a dedicated editing client computer. The video content is centrally hosted on storages that are based on common file systems and servers. The client is not directly integrated into the same local area network (LAN) as the content. Instead, it is connected via a wide area network (WAN) connection, utilizing a VPN. Files are accessed via common protocols such as SMB, AFP, and NFS. The editing solution can be used with HiRes video files or can work on proxy files, depending on the feature set of the application and configuration.

Advantages

- Allows simple scaling of editing clients, as long as the bandwidth of the VPN is sufficient.
- An editing application running on the client itself provides the full experience of the application, including controls, responsiveness, and file upload.

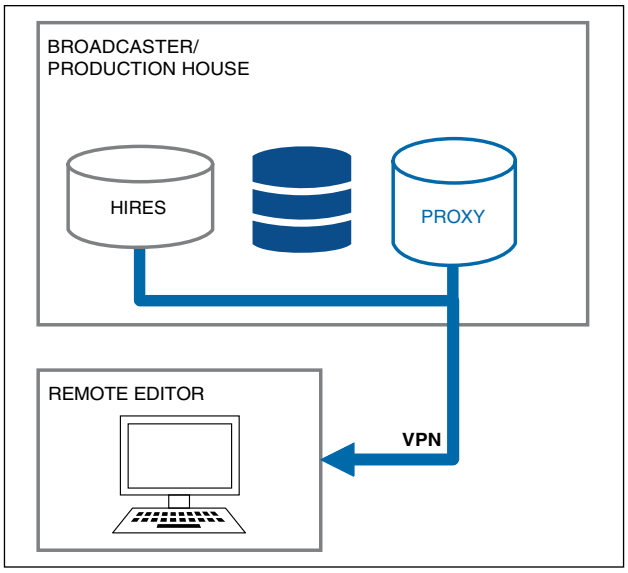


FIGURE 3. VPN.

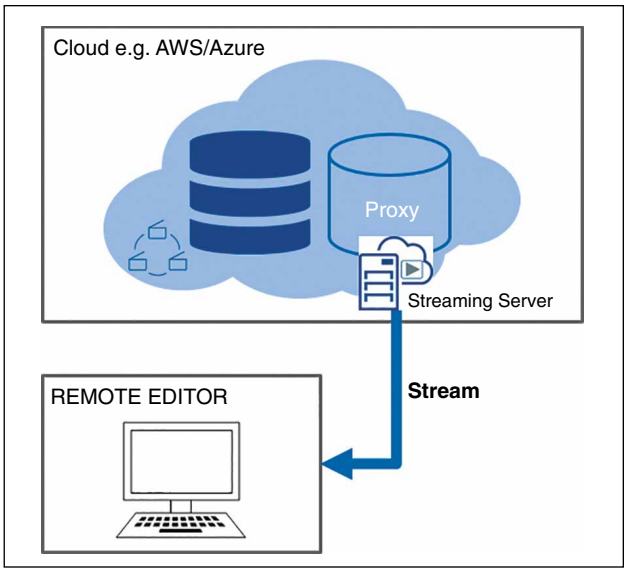


FIGURE 4. Cloud editing.

- Established collaboration workflows can be utilized.
- Content is protected via the encryption of the VPN tunnel.

Disadvantages

- Possible bandwidth limitations owing to VPN. Since the connection is routed over at least one VPN server, latency will be high.
- If proxy is used, the compression and codec influence the quality of the content. Some features like effects or focus approval might not be available.
- File access protocols are not designed for networks with high latency. SMB, for example, is block-based, and each time a block is transferred, the client communicates to the server. This results in a total latency that is at least equal to the latency between the client and the server.[8]

Cloud Editing—Proxy Streaming

This approach is cloud-based (**Fig. 4**). Hence, it runs the editing application on the client. So, the look and feel of the application are the same as with an existing, proven editing application. In this approach, the video content is stored on a public or private cloud storage. Instead of direct file access, the video is delivered to the client via a streaming server that streams the proxy file to the client application. Protocols used for streaming are Hypertext Transfer Protocol (HTTP) or Transmission Control Protocol (TCP). A streaming server allows—depending on the codec and bandwidth used—the delivery of dozens or hundreds of parallel streams. The streaming server needs to be known to the editing client and to request a dedicated video. The client needs to be aware of a reference that can be called on the video. Thus, a management layer like media asset management (MAM) is required.

Advantages

- The scaling of additional editing clients is done mainly at the client level. Additional streaming server instances can be implemented using cloud autoscaling functionalities.
- With the editing application running on the local client, the full functionality and experience of the client can be used.
- The local client enables local ingest via editing client functionalities.
- The use of proxy files reduces the required bandwidth; thus, the costs of cloud file transfers are reduced.
- Seamless transfer of clients working remotely and "in house." The client installation keeps its settings and additional plugins that can be used in both environments.

Disadvantages

- Owing to the usage of proxy, the quality of the content is not on a par with the original HiRes video.
- Bandwidth of the connection can be limited depending on the used access (Long-Term Evolution (LTE), fiber, etc.).
- An additional management layer is required to orchestrate the streaming server, whereas the management layer can provide additional functionalities enhancing the collaborative editing functionalities.

Cloud Editing with Live Encoding

An existing solution (**Fig. 5**) enables editing clients to handle video from a streaming source. It is based on a streaming engine that transcodes the source video based on the available bandwidth on the fly to the client. This means that each stream requires a dedicated compute resource to deliver video to the client.

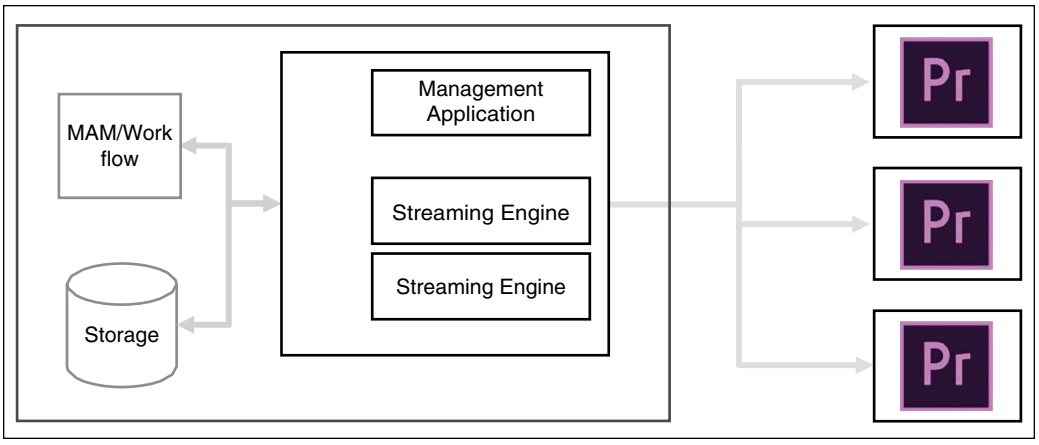


FIGURE 5. Cloud editing with live encoding.

Advantages

- Streaming provides the best possible quality depending on the available bandwidth.
- With the editing application running on the local client, the full functionality and experience of the client can be used.
- The local client enables local ingest via editing client functionalities.

Disadvantages

- Relatively high numbers of streaming servers to serve a typical number of editing clients.
- High costs driven by hardware requirements for live transcoding.
- Additionally, the streaming servers continuously require direct access to the source files, which limits the flexibility of the outlet of the video stream.
- In most cases, the streaming servers would be hosted on premise along with the production system. This limits the suitable location of reachable editing clients because, with growing distance, the latency of the network prevents smooth operations.

Derived Findings

Based on the previous examples and requirements, the following consequences can be derived.

Bandwidth: Since the bandwidth is limited, the production formats cannot be used directly for editing. Thus, a proxy format seems to be the appropriate approach. Hence, the quality needs to be as close to the original file as possible.

Availability: A cloud solution provides the ability to scale streaming servers based on the number of connected clients. Moreover, a public cloud solution hosted on major platforms like Microsoft Azure (MS Azure) or Amazon Web Services (AWS) allows the creation of streaming instances close to the client's location, resulting in lower latency.

Tooling: Integrating a streaming solution into a commonly used editing application like Adobe Premiere Pro allows editors to work with well-known software and does not require adjustment to new clients or workflows.

Security: Security is always an important aspect when dealing with IT systems. But when it comes to the cloud, the setup gets more vulnerable, since servers and storages are technically reachable via the internet. Even though this is a large field with several aspects, big cloud providers offer solutions to protect data and put effort into securing connections by providing detailed guidelines and restrictions[9] on how to secure systems, which are hosted in the cloud. Although the providers guide the process to create security, they still require businesses to be responsible and to apply the necessary measures.

On top of that, the transported data needs to be secure so as to provide privacy and data integrity. Encryption algorithms, which are used in the Transport Layer Security [TLS, formerly known as *Secure Sockets Layer* (SSL)], prevent third parties from reading and modifying any transferred information. Encryption needs to be applied to secure sensitive information or content that is regulated by copyright.[10]

Cloud Editing based on a Media or Production Asset Management System

Another approach to remote editing is based on the same technology as the described cloud editing along with live encoding. It uses streaming servers and compressed video for playback in the editing client, but on top of that, resources of a MAM system are utilized. A MAM system can provide components that can build a base for remote editing in the cloud. Associated with that are pregenerated proxy video, metadata enrichment, and management of editing projects.

Technical Overview

An example of remote editing (**Fig. 6**) in the cloud that is built on top of a MAM system is as follows. It shows an overview of a hybrid craft editing installation, extended with remote editing.

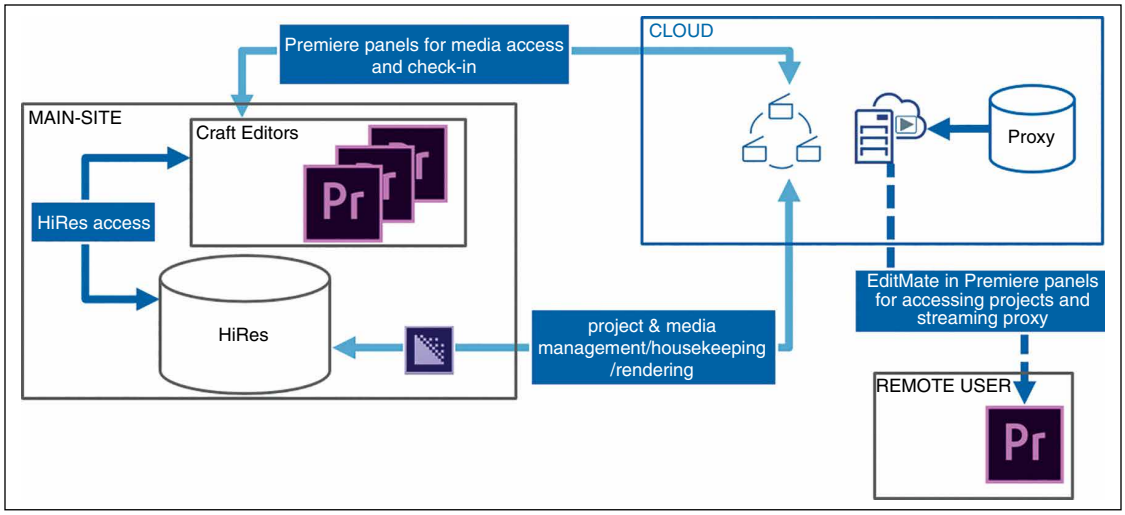


FIGURE 6. Cloud editing and hybrid installation.

The main site contains a "classic" setup of on-premise-based craft editing. Media assets are centrally stored on a HiRes storage. These files are accessed by local Adobe Premiere installations. The craft editor imports the video assets to its bin, edits a sequence and renders it via the local Adobe Media Encoder, to create a new asset. All benefits and limitations of an on-premise solution remain unchanged.

On top of that, there is a cloud solution extending the range of editing functionalities to remote locations. A project and media management solution is hosted in the cloud, which enables the local and remote editor to search and browse for centrally managed assets. The managed assets can be stored in the cloud or in the on-premise storage. The proxy, created off the HiRes source files, is also located in the cloud, which also applies to the streaming server. The streaming server accesses the proxy and streams it to the connected remote clients.

Although the architecture might vary depending on the system scaling, a solution hosting HiRes and renderer in a cloud environment would be viable.

Proxy Format

Selection of the proxy format has a significant influence on the perception of the edit experience since resolution and compression influence the availability to determine and evaluate quality or sharpness. Hence, improved quality results in higher bitrate requirements and therefore processing power or processing time to create the proxy file.

A proven proxy format standard is SMPTE RDD25. It is an Advanced Video Coding (AVC) "Long GOP" proxy with AAC audio, originally conceived to standardize low-resolution proxies for use with low-res editors. With standard definition (SD) resolution (640 × 360) and a bitrate up to 2 Mb/s and four stereo audio channels, it allows fast encoding and decoding, but lacks the ability to provide a HiRes experience (**Fig. 7**) to the user. Thus, the existing format was extended in terms of resolution, bandwidth and audio channels. Based on the HiRes source and using the main instead of the base profile, a proxy was created, which can encode faster than realtime (depending on the source file up to 70 frames/s). H.264 with the speed of 6–10 Mb/s resolution of 1920 × 1080 and eight stereo audio tracks is getting closer to the source video, but meets with its parameters the requirements of limited bandwidth. The proxy can be created in an mp4 container, which extends the interoperability or in a Material Exchange Format (MXF) container that allows editing while the proxy is generated. The following illustration shows a comparison of details between the original XDCamHD HiRes (**Fig. 8**) and the H.264 10 Mb/s proxy (**Fig. 9**). Fine details of the structure on the roof and windows show compression artifacts, but the overall perceived quality of the proxy comes closer to the original; see "Analysis of Look and Feel."

With evolving compression technology, new codecs and container formats will extend the possibilities of proxy editing by allowing better quality with a similar frame rate or vice versa. High-efficiency video coding (HEVC) also known as H.265 is a compression standard that offers about two times the data compression ratio in comparison to AVC-like H.264.[11] It is not yet widely distributed. However, it is becoming more and more popular because earlier limitations like high demand for processing power have become less critical. Newer generations of central processing units (CPUs) include dedicated HEVC decoding and allow more efficient playback. Hence, it is a good candidate to replace H.264 for this use case. However, it still requires a certain amount of licensing costs to use it. Open royalty-free video codecs such as AO-Media Video 1 (AV 1) and VP9 offer an alternative. With VP9, Google created a

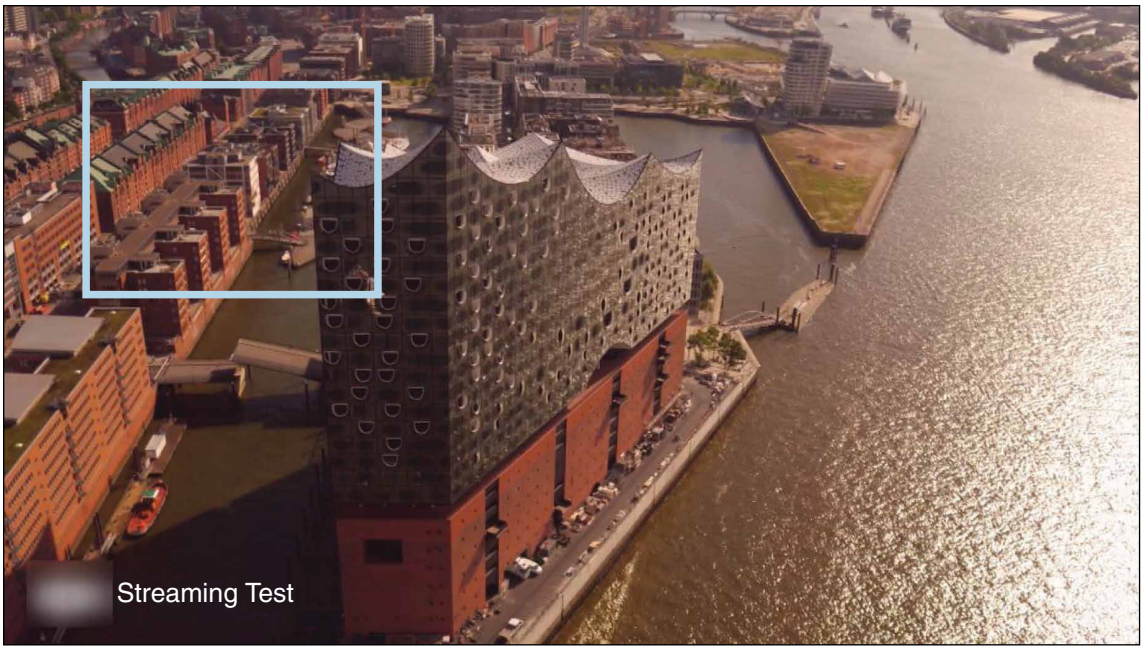


FIGURE 7. Source video.

competing codec, which is becoming more widespread, and used on the world's largest video distribution platform YouTube. With AV 1 and the recent start of its beta test on YouTube,[12] the battle between new codecs and technologies entered the next round.[13]

Peak Files

A peak file is a wave form representing an audio level in a graph based on the time axis. Adobe Premiere creates this wave form on every object within the editing sequence. It analyzes the file and renders the view. Depending on the length of the imported content, the process takes from seconds up to several minutes. During the rendering process, the editing client is slowed down and the editing usability is heavily compromised. For an editor, this is time lost. An MAM as a basis for the proxy editing can create these peak files already during ingest. Adobe Premiere can import these prerendered wave forms, which reduces the analysis of files on each client. Prerendered peak files, which are available immediately after import, speed up the workflow.

Latency and Perceived Response

The streaming protocol is a fundamental element for cloud editing. The subjectively perceived experience of editing stands and falls with the performance of playback and responsiveness, which is mainly driven by the performance of streaming. A frequently used streaming protocol is MPEG DASH,[14] which is used by the big online streaming platforms YouTube and Netflix[15] in the HTML5 context. The requirement of these platforms is mainly to provide smooth linear forward playback. Although seeking is possible, scrubbing suffers from large segments that need to be transferred from the server to the client. To simulate a fast scrubbing functionality, YouTube and Netflix use a "trick mode" that is based

FIGURE 8. XDCamHD 50 Mb/s.

FIGURE 9. H.264 10 Mb/s.

on thumbnails. The user navigates via thumbnails to the designated position. A protocol is needed that meets the required perception and feel of an editing client directly accessing the HiRes files.

Adobe Premiere allows the integration of custom-made proprietary importer plugins that handle the playback of the content. The content can be stored locally on a network storage or provided via a streaming server. A proprietary protocol has been introduced by Arvato that utilizes a TCP connection. This article describes the overall functionality and improvements, but will not explain the exact details of the implementation owing to the protection of intellectual property.

The implemented Premiere Importer functionality has been adjusted to transport only the exact required individual frames as they are requested by the client application. This allows fast scrubbing as well as fast forward and playback. A video may be divided into many files (called chunks or segments), each containing only a few seconds of video at one extreme or stored in a single unchunked file at the other.[16] With larger chunks, it might happen that a single frame is requested, but two complete chunks are transported and decoded, because the frame is within a Group of Picture (GOP) that is spread over two chunks.

Producers—especially in use cases of sports and news—frequently need to scrub through large amounts of video to find the elements they need for their projects. When only the frames needed for decoding are downloaded, latency in streaming is reduced, leading to a better user experience. Improvements in latency optimize the workflow for editors and producers. Subjective tests with operators and editors show that a latency of >30 ms results in poorly perceived experience of editing. So different measures need to be applied to keep the latency low.

The asynchronous sending and receiving of packages would be an improvement. Asynchronous frame requests improve response times. In an asynchronous frame request scenario, the client can send multiple requests at the same time while, in parallel, receiving all return information.

Latency depends on the quality of the network and especially on the distance between the streaming server and the client. Therefore, the outlet of a cloud environment needs to be as close to the client as possible. The large hyper scalers AWS and MS Azure enable this with their distribution of data centers across the globe.

Initially, the main purpose for streaming servers was to channel access to the source files by providing video upon request and preventing direct access to it. This avoids inadvertently moving, copying, or deleting files or other unwanted action. With the introduction of low bandwidth and long network latency owing to Cloud, WLAN, or VPN, an additional functionality became more relevant: smart random access to the file.

MXF as a quasi-standard container for video in a broadcast context enables editing on growing files. This comes along with a random index pack and an index table, which needs to be opened and searched for. This results in several small read operations. SMB via an on-prem network connection is a suitable protocol. But SMB is not really designed to operate via high-ping networks, resulting in high potential for packet loss, which results in degraded performance.

Common storage types that are used by AWS and Microsoft for media files are object storages such as S3 and Blob. Object storages manage data as objects, as opposed to other storage-like file systems (that manage data as file hierarchy) and block storages (that manage data as blocks within sectors and tracks). Object storage provides some advantages compared with file storage, like better performance on big content and throughput. Data can be stored across multiple regions, scaling infinitely to petabytes and beyond. The objects can be enriched with metadata.[17] But when it comes to cloud editing, some obstacles need to be overcome.

If an editor imports an MXF file directly into Premiere from a blob storage, the performance of the importer would be suboptimal, because the MXF parser needs fast random access to be able to work smoothly. However, this is possible with object storages—even though Azure Blob Storage offers options for fast random access, this cannot be used for growing file support. The applied streaming technology solves this requirement for the storage. The client does not need fast random access to the server: frames are only transferred to the exact byte. The requirements for random access are the server side in this case. But how can we connect cloud storage to a streaming server with sufficient random access performance? At the time being, this is done by locating the streaming server and storage in the same availability zone.

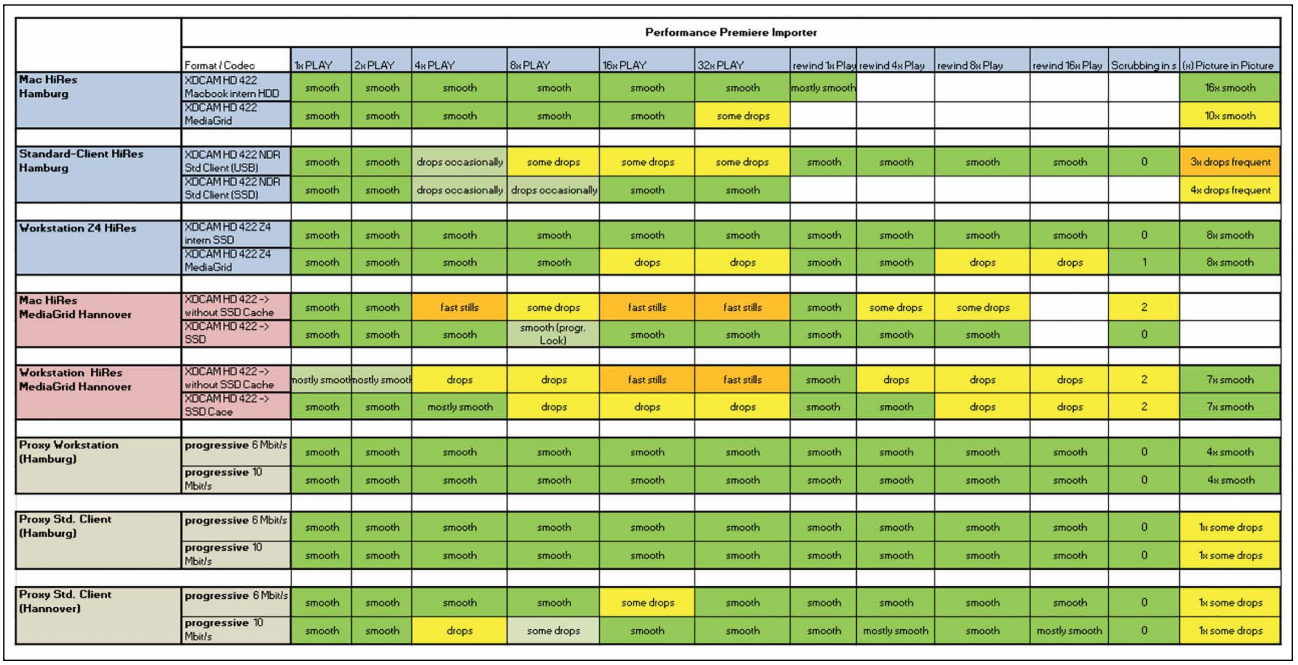
Analysis of Look and Feel

Evaluation of Performance

An analysis (**Fig. 10**) of a German public broadcaster has been done that examines the solution as described above. This analysis compares different client computers and format scenarios. Experienced editors and engineers provided a subjective rating of the look and feel of the editing performance. The proxy files, streaming server, and client were located on premise. The editing clients were connected with a 1 Gb/s network connection to the streaming servers.

The goal of this analysis was to figure out the possibility of combining four different editing sites into a single installation. The editing sites are allocated at different sites, hundreds of kilometers apart.

The existing HiRes craft editing installation was based on Microsoft Windows 10 and Apple MacOS



	Performance Premiere Importer												
	Format / Codec	1x PLAY	2x PLAY	4x PLAY	8x PLAY	16x PLAY	32x PLAY	rewind 1x Play	rewind 4x Play	rewind 8x Play	rewind 16x Play	Scrubbing in s	(s) Picture in Picture
Mac HiRes Hamburg	XDCAM HD 422 Macbook intern HDD	smooth	smooth	smooth	smooth	smooth	smooth	mostly smooth					16x smooth
	XDCAM HD 422 MediaGrid	smooth	smooth	smooth	smooth	smooth	some drops						10x smooth
Standard-Client HiRes Hamburg	XDCAM HD 422 NDR Std Client (USB)	smooth	smooth	drops occasionally	some drops	some drops	some drops	smooth	smooth	smooth	smooth	0	3x drops frequent
	XDCAM HD 422 NDR Std Client (SSD)	smooth	smooth	drops occasionally	drops occasionally	smooth	smooth						4x drops frequent
Workstation Z4 HiRes	XDCAM HD 422 Z4 intern SSD	smooth	smooth	smooth	smooth	smooth	smooth	smooth	smooth	smooth	smooth	0	8x smooth
	XDCAM HD 422 Z4 MediaGrid	smooth	smooth	smooth	smooth	drops	drops	smooth	smooth	drops	drops	1	8x smooth
Mac HiRes MediaGrid Hannover	XDCAM HD 422 -> without SSD Cache	smooth	smooth	fast stills	some drops	fast stills	fast stills	smooth	some drops	some drops		2	
	XDCAM HD 422 -> SSD	smooth	smooth	smooth	smooth (progr. Look)	smooth	smooth	smooth	smooth	smooth		0	
Workstation HiRes MediaGrid Hannover	XDCAM HD 422 -> without SSD Cache	mostly smooth	mostly smooth	drops	drops	fast stills	fast stills	smooth	drops	drops	drops	2	7x smooth
	XDCAM HD 422 -> SSD Cace	smooth	smooth	mostly smooth	drops	drops	drops	smooth	smooth	drops	drops	2	7x smooth
Proxy Workstation (Hamburg)	progressive 6 Mbit/s	smooth	smooth	smooth	smooth	smooth	smooth	smooth	smooth	smooth	smooth	0	4x smooth
	progressive 10 Mbit/s	smooth	smooth	smooth	smooth	smooth	smooth	smooth	smooth	smooth	smooth	0	4x smooth
Proxy Std. Client (Hamburg)	progressive 6 Mbit/s	smooth	smooth	smooth	smooth	smooth	smooth	smooth	smooth	smooth	smooth	0	1x some drops
	progressive 10 Mbit/s	smooth	smooth	smooth	smooth	smooth	smooth	smooth	smooth	smooth	smooth	0	1x some drops
Proxy Std. Client (Hannover)	progressive 6 Mbit/s	smooth	smooth	smooth	smooth	some drops	smooth	smooth	smooth	smooth	smooth	0	1x some drops
	progressive 10 Mbit/s	smooth	smooth	drops	some drops	smooth	smooth	smooth	mostly smooth	smooth	mostly smooth	0	1x some drops

FIGURE 10. Performance evaluation.

clients connected to a central Harmonic MediaGrid Server storage. Each client had an available bandwidth of 1 Gb/s.

The test cases were as follows.

- Playback with different speeds (**Fig. 11**) (1×, 2×, 4×, 8×, 16×).
- Rewind playback with different speeds (1×, 4×, 8×, 16×).
- Scrubbing and navigating directly to different positions on the timeline.
- Playback of picture in picture.

The blue marked fields represent the HiRes editing setup in the city of Hamburg. Three different server types were tested with internal hard disk drive (HDD) and central access to the MediaGrid.

The red marked fields represent another remote access via the city of Hannover where the clients were MAC and an HP Z4 Workstation.

The gray field represents the results of HD-proxy editing. The proxy had been created with 6 and 10 Mb/s.

Two cases compared in detail underline the performance and possibilities of the technology used.

The streaming server was located in Hamburg, Germany, and the clients were either in Hamburg or Hannover. Hamburg and Hannover are connected via a 10 Gb/s connection, which is used for all network traffic between these two entities.

Workstation Hires and Workstation HD Proxy Streaming

A typical workstation (**Fig. 12**) for applications that require medium performance has been used as a craft editing client at the German broadcaster that ran the tests and evaluation.

The reference test, which is in all test scenarios evaluated as smooth, has been performed in an ideal scenario, where the HiRes source files were locally available on that client using an solid state drive (SSD) Hard Drive.

The typical scenario for production at this broadcaster is a setup of the second test on the Z4, which is used for daily news, documentary, and feature production. The files are stored centrally as house format XDCamHD on the Harmonic MediaGrid Production Storage. This results in access and transfer via the on-premise network.

Compared to the first series of tests, a slight decline of performance is noted, since fast forward and backward play results in notable drops. Hence, it does not influence its daily usage.

Workstation Z4 HiRes	XDCAM HD 422 Z4 intern SSD	smooth	smooth	smooth	smooth	smooth	smooth	smooth	smooth	smooth	smooth	0	8x smooth
	XDCAM HD 422 Z4 MediaGrid	smooth	smooth	smooth	smooth	drops	drops	smooth	smooth	drops	drops	1	8x smooth

FIGURE 11. Workstation with SSD and MediaGrid access test series.

Proxy Workstation (Hamburg)	progressive 6 Mbit/s	smooth	smooth	smooth	smooth	smooth	smooth	smooth	smooth	smooth	smooth	0	4x smooth
	progressive 10 Mbit/s	smooth	smooth	smooth	smooth	smooth	smooth	smooth	smooth	smooth	smooth	0	4x smooth

FIGURE 12. Workstation—HD proxy test series.

The same client was used, within the same location, to evaluate the performance with streamed HD proxy. The result was in all cases smooth and closer to the reference as the production setup.

Standard Client HiRes and Standard Client HD Proxy

A standard client (**Figs. 13–15**) at the broadcaster is a typical office PC that does not necessarily provide hardware for craft video editing. Therefore, the results for

Standard-Client HiRes Hamburg	XDCAM HD 422 NDR Std Client (USB)	smooth	smooth	drops occasionally	some drops	some drops	some drops	smooth	smooth	smooth	smooth	0	3x drops frequent
	XDCAM HD 422 NDR Std Client (SSD)	smooth	smooth	drops occasionally	drops occasionally	smooth	smooth						4x drops frequent

FIGURE 13. Standard client HiRes test series.

Proxy Std. Client (Hamburg)	progressive 6 Mbit/s	smooth	smooth	smooth	smooth	smooth	smooth	smooth	smooth	smooth	smooth	0	1x some drops
	progressive 10 Mbit/s	smooth	smooth	smooth	smooth	smooth	smooth	smooth	smooth	smooth	smooth	0	1x some drops
Proxy Std. Client (Hannover)	progressive 6 Mbit/s	smooth	smooth	smooth	smooth	some drops	smooth	smooth	smooth	smooth	smooth	0	1x some drops
	progressive 10 Mbit/s	smooth	smooth	drops	some drops	smooth	smooth	smooth	mostly smooth	smooth	mostly smooth	0	1x some drops

FIGURE 14. HD proxy test series.

	Requirement	Quality Evaluation			Comment
		2 Mb/s	6 Mb/s	10 Mb/s	
Multicam	Multicam Edit 2 sources parallel	×	✓	✓	Simple studio production
	Multicam Edit 6 sources parallel	×	✓⇗	✓⇗	e.g. concert with more than 5 cameras
	Multicam Edit more > 6 sources	×	✓⇗	✓⇗	e.g. complex concert production with more than 6 streams
Titling	ProRes titles	✓	✓	✓	Titles are imported to project as HiRes graphics
	Titling with easy insert	✓	✓	✓	Titles are imported to project as HiRes graphics
	Motion graphics templates (mogrt)	✓	✓	✓	
	Delivery master: FCC + Mix+ Inserts	✓⇗	✓⇗	✓⇗	
	Set and edit subtitles	✓	✓	✓	
Rough Edit Workflows	Browse material	✓	✓	✓	
	Collect content for project	✓	✓	✓	
	Simple edit with rough cuts	✓	✓	✓	
	Nonlinear studio edit	×	✓	✓	Image sharpness needs to be evaluated
	Rough-cut for program	✓	✓	✓	
	Trimming	✓	✓	✓	
	Assemble archive content	✓	✓	✓	
	Voice Over	✓	✓	✓	
Approval	Editorial acceptance, internal control	✓	✓	✓	
	Technical quality control	×	×	×	
Material	Live server cut (feed ingest) growing files	✓	✓	✓	Proxy is generated on growing file
	Take over VJ projects ("refine")	×	×	×	Depending on project size, the proxy transcoding required
	Evaluation of pixelated material (artifacts) (LiveU)	×⇗	×⇗	×⇗	Artifacts from the source cannot be distinguished from the artifacts of the stream
	Working with drone material (origin: 4K)	✓⇗	✓⇗	✓⇗	Downscaling causes moiré patterns and line flicker
	Edit with archive material	✓⇗	✓	✓	e.g. differentiate SD and HD source
	Aspect ratio recognizable (4:3, 16:9, LB, PB, zoom)	✓	✓	✓	For archive content of false scaled content
	Native cut with S-log, C-log, 10-bit material	×	×	×	Working with lookup tables not possible. Proxy does not provide 10-bit dynamic range

Continued

Effects	Cut with simple clip or transition effects	✓⇗	✓	✓	Wipe, white/black flash, transition etc.
	Cut with elaborate effects (multilayering)	×	✓⇗	✓⇗	Effects on multiple layers, key framing etc.
	Cut with effects like C. pinning or split screens	×	✓	✓	C. pinning or split screens
	Cut with after effects elements	×	×	×	After effects projects with dynamic link relation
	Animate photos, Ken Burns effect etc.	×	✓⇗	✓⇗	Key frame editing, Bezier curves. etc.
	Simple pixelations	✓	✓	✓	
	Pixelations and (auto) tracking	×	✓⇗	✓⇗	
	Simple keys	×	×	✓⇗	Assessing clean key is difficult, because artifacts might be result of compression or key
	Elaborate Keys	×	×	×	Exact settings of the keys with spill and color suppression etc.
	Working with masks	×	✓	✓	
	Edit with graphical templates	✓	✓	✓	
	Edit with with trailer packaging	×	×	×	After effects integration
	Slow-motion or time-lapse	✓⇗	✓⇗	✓⇗	
	Slow-motion curves and speed ramps	✓⇗	✓⇗	✓⇗	
Special Cases	Create web videos (div. aspect ratio, etc.) without after effects	×	✓	✓	
	360° Videos	×	×	×	Including stitching - requires dedicated plugins and software
	Simple color correction on the editing client	×	✓	✓	Standard pc lacks calibrated monitors
	Elaborate color correction, color matching	×	×	×	Would theoretically be possible in premiere (HiRes) - but is not performed
	Highlight edit incl. Playback from timeline during live events	×	✓⇗	✓	Sports - halftime edit etc.
	Edit with short time to air	✓	✓	✓	Proxy is generated on growing file
	Projects with various distribution targets	✓	✓	✓	

FIGURE 15. Quality evaluation.

the test series show declined performance. Since these clients do not have access to the production storage, the test series was performed with the local hard drive and a mobile USB hard drive.

The test series with HD proxy shows a better performance than the direct HiRes access.

The performance within the entity Hamburg shows a smooth playback both for 6- and 10-Mb/s HD proxies, whereas the decoding of picture in picture results in some drops. Also, the playback in the remote entity Hannover shows acceptable performance, with some additional drops.

With HD proxy, it seems possible to use existing office clients for editing when needed.

The overall results show that depending on the client type used and the connection to the storage, the performance varies even in a HiRes editing case. Overall, HiRes and HDProxy editing shows a similar performance, which indicates that HDProxy editing is a viable supplement for existing HiRes editing solutions. Except for the picture-in-picture mode with multiple streams, the playback was described as "smooth" with some dropped playback in fast forward.

Evaluation of Quality

The German broadcaster (**Fig. 15**) performed additional tests to evaluate the quality of the utilized proxy. Experienced editors rated the subjective quality of the

video, categorized typical use cases and evaluated the feasibility of those.

A test series was conducted with 2 (SD), 6 (HD), and 10 (HD) Mb/s H.264 HD proxy. Again, a reference HP Z4 workstation and an office client served as the basis. A red x indicates cases in which the quality of the proxy was not sufficient. A light green arrow signifies cases that can be performed with proxy editing, but would require a final approval, ideally on the used HiRes or rendered clip. The dark green fields have been approved as cases which can be realized with proxy editing only. It is obvious that several cases cannot be done with 2 Mb/s SD proxy, since the resolution does not allow the evaluation of sharpness, focal length or other technical parameters. Based on the parameters, a proxy will not match the quality of the original HiRes file, but utilizing an H.264 video codec with full HD resolution and a bit rate between 6 and 10 Mb/s, results can be achieved that are acceptable to editors without significant detriment to the user experience.

HD proxy shows good results on "standard" edits such as titling, rough editing, simple effects, graphical templates, and multicamera productions, whereas elaborate editing, which requires technical evaluation and sophisticated graphics, pushes the limits of proxy editing. For those cases, a viable setup could be a combination of HD proxy editing, with proxy hosted in the cloud and HiRes editing on premise. Several edits can already be done and finished remotely, but when required, those edits can be completed on premise or exchanged with editors, which have access to the HiRes files.

An additional test series (**Fig. 16**) was performed during the Arvato User Group Meeting 2018. Different test clips were shown to an audience of ~120 experienced broadcast experts and engineers. The test clips showing a Siemens star and bars could not clearly be determined as proxy or original XDCAM HD, whereas a sequence of a drone flight was more obvious. The sequence of the drone flight did show more artifacts on details and areas of low structure. Nevertheless, engineers evaluated the test installation as suitable for producers and craft editing.

The poll question "For which user groups could HD remote editing be a replacement or extension to the current craft editing?" provided these results.

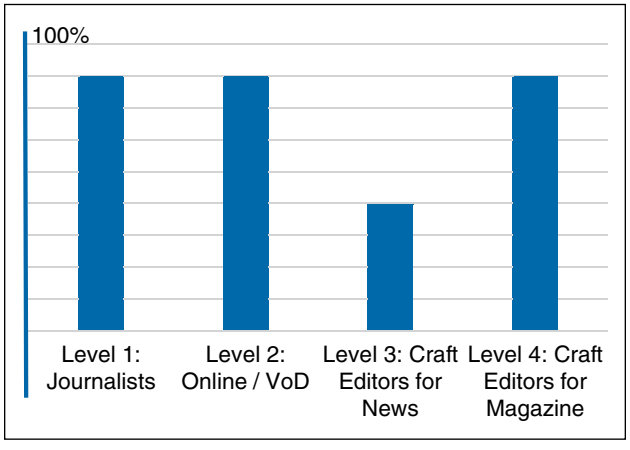


FIGURE 16. HD remote editing user group evaluation.

Conclusion

For a runner, for example, Usain Bolt, to improve his speed, there is no single element that he works on. He and his team work on improving multiple elements of his performance: if his stride can be a centimeter longer, his starting time a quarter of a second faster, then he will achieve a still faster 100 m. In the same way, there is no single change that will improve and accelerate cloud-based editing workflows to the extent that cloud-based editing becomes the norm.

What is needed are teams of experts with an in-depth understanding of various elements to work on each area—server–storage connectivity, the use of TCP, reduction in the number of frames transferred, improved streaming protocols and better handling of audio files, as well as smart local caching.

Cloud editing is possible. With these improvements, it can become a reality.

References

1. The 2018 Devoncroft Global Media Technology Trend Index.
2. Microsoft Technet.
3. Gartner Glossary.
4. "DD 25:2014—SMPTE Registered Disclosure Doc—AVC MXF Proxies." [Online]. Available: https://ieeexplore.ieee.org/document/7290564/
5. "The Art of Film and Video Editing Part-1." [Online]. Available: https://www.videouniversity.com/articles/the-art-offilm-and-video-editing-part-1/
6. "Average Transmission Speed of Internet Connections in Germany from 3rd Quarter 2007 to 1st Quarter 2017." [Online]. Available: https://www.statista.com/statistics/480570/broadband-average-internet-speed-germany/
7. "Protecting Netflix Viewing Privacy at Scale." [Online]. Available: https://medium.com/netflix-techblog/protectingnetflix-viewing-privacy-at-scale-39c675d88f45
8. "SMB/CIFS Performance over WAN Links." [Online]. Available: https://blogs.technet.microsoft.com/neilcar/2004/10/26/smbcifs-performance-over-wan-links/
9. "AWS_Security_Best_Practices." [Online]. Available: https://d0.awsstatic.com/whitepapers/Security/AWS_Security_Best_Practices.pdf
10. "Protecting Netflix Viewing Privacy at Scale." [Online]. Available: https://medium.com/netflix-techblog/protectingnetflix-viewing-privacy-at-scale-39c675d88f45
11. "H.265/HEVC vs H.264/AVC: 50% Bit Rate Savings Verified." [Online]. Available: https://www.bbc.co.uk/rd/blog/2016-01-h-dot-265-slash-hevc-vs-h-dot-264-slash-avc-50-percent-bit-rate-savings-verified
12. "YouTube Begins Beta-Testing AV1 CODEC on Beta Web-Browsers." [Online]. Available: https://www.techpowerup.com/247584/youtube-begins-beta-testing-av1-codec-on-beta-web-brows-ers
13. "Codec Wars: The Battle between HEVC and AV1." [Online]. Available: https://www.ibc.org/delivery/codec-warsthe-battle-between-hevc-and-av1/2710.article
14. "MPEG-DASH an Overview." [Online]. Available: https://www.encoding.com/mpeg-dash/

15. "Why YouTube & Netflix Use MPEG-DASH in HTML5." [Online]. Available: https://bitmovin.com/status-mpeg-dash-today-youtube-netflix-use-html5-beyond/
16. "To Chunk or not to Chunk: Implications for HTTP Streaming Video Server Performance." [Online]. Available: https://cs.uwaterloo.ca/~bernard/nossd47.pdf
17. "Object Storage vs. File Storage: What's the Difference?" [Online]. Available: https://cloudian.com/blog/objectstorage-vs-file-storage/

About the Authors

Ulrich Ening is a broadcast and IT specialist with extensive experience in the broadcast and software industry. He has more than ten years of experience in media asset management technology, interfacing with a large number of components within the broadcast workflow. He is an engineer and solution-architect with Arvato Systems S4M, Koeln, Germany, and is working on various projects in Europe, North America, and Middle East. Since the end of 2019, he has been part of the Video Production Management Suite (VPMS) product management team at Arvato Systems, Cologne, Germany.

Karsten Schragmann studied media technology at the Cologne University of Applied Sciences, Cologne, Germany, while also working as a freelance event technician at international corporate events and trade fairs. After working in product management at Sedus Stoll AG, Hattershei, Germany, and Riedel Communications, Wuppertal, Germany, in 2014, he joined Arvato Systems.

Schragmann was initially in charge of managing a range of projects and products. Since June 2018, he has been responsible for strategic positioning and further development of the production asset management solution VPMS EditMate.

Presented at the SMPTE 2018 Annual Technical Conference & Exhibition, Los Angeles, CA, 22–25 October 2018.