F R O S T & S U L L I V A N

Virtual Workflows Have Strong Impact on Modern Video Architectures:
New Applications Creating Growth Opportunities in OTT Viewing

A Frost & Sullivan White Paper

Sponsored by
Hewlett Packard Enterprise

CONTENTS

TRENDS IN VIDEO CONSUMPTION AND IMPLICATIONS ON WORKFLOW

On-demand, user-driven content consumption is overtaking conventional broadcast workflows. Industry trends and metrics emphatically underscore this shift. Global Over-the-Top (OTT) revenues, high as they already are, are expected to more than double over the next five years.[1] Subscription VOD revenues are projected to outgrow advertising revenues, and growth forecasts for OTT subscribers continue to be energetically optimistic worldwide, even as conventional pay TV subscriber rates continue to trend downward. Pay TV operators and broadcasters are relying more and more on OTT for viewers and advertising revenue. Broadcasters are attracting comparable views across traditional and OTT viewing. New applications such as virtual reality and e-sports set the stage for even faster increases in video processing needs and video traffic. Cisco Systems predicts that by 2020[2], over 80% of data on broadband and mobile networks will be video.

Implications of the Breakdown of TV-by-appointment

The primary implication of the breakdown of TV-by-appointment and the rise of on-demand models is that fixed broadcast workflows are giving way to unicast-centric workflows. The quest to deliver managed content experiences on unmanaged devices requires extensive use of server-side video processing and middleware logic to curate, process, deliver, measure, analyze and monetize content experiences. The challenge of serving personalized content streams at high quality without onerous levels of infrastructural investment is a significant one.

The broadcasters and programmers who thrive and grow amidst this changing environment are those who can achieve agility and flexibility to quickly adapt to changing programming and formatting needs. Workflow changes today can take months or weeks; services must be able to bring this down to hours or minutes in order to remain competitive and responsive to user demand. This requires content businesses to transition away from hardware-centric, fixed-function infrastructure to open, IP-based, and software-defined workflows. A modern approach to transcoding and content delivery infrastructures is also necessary to simultaneously achieve scalability, cost-efficiency and high quality.

ARCHITECTURAL IMPERATIVE FOR CONTENT BUSINESSES: SOFTWARE-BASED WORKFLOWS DELIVER MUCH-NEEDED AGILITY

Managed experiences have traditionally been delivered over reliable broadcast infrastructure or via managed networks to set-top boxes. The crucial business imperative for video service operators today is to deliver the equivalent of these managed experiences to unmanaged devices over unmanaged networks, but still with consistent quality of experience. Broadcasters are increasingly disintermediating Pay TV providers to go directly to consumers via websites, apps and streaming device channels. Networks are facing a significant transition as they begin to market individual shows through their own apps or via services like Netflix or Amazon. One implication of this transition is that channel brands are becoming secondary to direct audience interest in specific shows or events.

The Only Constant is Change

OTT delivery to unmanaged devices brings complications to compression, delivery and rendering technologies and standards. These technologies can change quickly and unpredictably. Content workflows must consequently be able to rapidly adapt to ongoing changes, some examples of which include:

1 http://www.digitaltveurope.net/383162/netflix-to-help-ott-revenues-almost-double-in-next-five-years/
2 http://www.cisco.com/c/dam/en/us/solutions/collateral/service-provider/visual-networking-index-vni/complete-white-paper-c11-481360.pdf

- Compression and delivery formats are in flux. The most disruptive of these are the rise of HEVC as the potential successor to AVC; the growing adoption of DASH, even as HLS endures but SmoothStreaming fades; the transition of resolutions from sub-SD, past SD to full HD and now UltraHD; and emerging modalities such as immersive video and high dynamic range (HDR)/wide color gamut content. Workflows must be able to quickly adapt to changes in incoming source format, outgoing resolutions, compression and packaging technologies, and more.

- Broadcast regulations are increasingly applicable to live OTT transmissions. For example, live OTT content must comply with accessibility requirements for closed captioning and CALM regulations for loudness of advertising. As new regulations emerge with regional or country-wide impact, it must be possible to augment workflows with additional features or adjusted parameters without the need for significant re-engineering.

- Monetization technologies continue to evolve as businesses continue to experiment with pricing and tracking, and as analytics and recommendation technology become more intelligent and capable. For example, recent increases in advertisement blocking from mobile phones and quality-of-service considerations are driving a shift from player-side ad insertion to server-side ad insertion. When a workflow and rendering architecture is well-designed, these types of changes can be absorbed at relatively lower cost with higher speed and lower risk.

Embracing Change While Protecting the Bottom Line

As video services become increasingly consumer-driven, and as the consumer device ecosystem continues to get more fragmented, the above business imperatives will become increasingly critical. At the same time, business profitability considerations require that CAPEX be controlled and OPEX be optimized. In other words, quality and flexibility of services and scalability and agility of workflows must be achieved in tandem with a viable cost structure. As the technology landscape remains firmly in flux, it is clear, for example, that fixed-function equipment can no longer be expected to remain relevant for its useful lifetime. The only way to protect CAPEX today is to invest in flexible infrastructure powered by software. Traditionally, this forced a tradeoff between flexibility, performance and density. Fortunately, advances in general-purpose processors like the Intel® Xeon® Processor E3-1500 v5 Product Family[3] with Iris™ Pro Graphics are changing the status quo and pulling up software-based systems to perform at comparable levels to dedicated hardware systems, but without the risk of early obsolescence. These modern chipsets simultaneously deliver (a) the power efficiency, acceleration and processing density traditionally only available with dedicated hardware, and (b) the flexibility of software on general-purpose processors. These processors provide built-in hardware acceleration for broadly used AVC as well as a new HEVC standard that supports best-in-class quality and throughput. Rich interfaces allow individual vendors to layer in specialized IP to provide differentiated media processing capability. At the same time, because servers such as the HPE Moonshot system (with Intel® Xeon® Processors inside) are supported by nearly all vendors, customers benefit from the option to choose components for their workflow from a wide variety of competitive options.

COUNTERING THE HARDWARE MYTH: SOFTWARE VIDEO PROCESSING CAN EFFECTIVELY POWER CARRIER-GRADE WORKFLOWS

Conventional wisdom often dictates that software-based workflows cannot deliver the same levels of density, power efficiency, compression quality, latency and reliability as hardware-based infrastructure. With modern solutions like the Intel Xeon processor-based HPE Moonshot system, however, the equation has shifted significantly in favor of a win-win situation. Technology decision makers should fully familiarize themselves with the capabilities

3 This processor was formerly known as Skylake

of state-of-the-art systems and make decisions based on real capabilities rather than perceived risk. In this section we summarize findings from real-world tests that demonstrate that best-in-class servers can deliver competitive performance, while also delivering the flexibility and repurpose-ability that hardware alone cannot provide.

Cost

The exact cost per channel in terms of CAPEX varies based on factors such as video resolution in terms of pixels, color depth and frame rate; compression use case in terms of encoding versus transcoding and file versus live; and actual codec. In general, we find that the cost per channel for software running on commercial off-the-shelf (COTS) hardware is comparable to hardware alternatives. For example, channel prices for transcoding of full HD streams are in the range of one to a few hundred US dollars per stream. This use case accounts for the high volume of deployments. In the case of full HD real-time contribution encoding in AVC, which represents the higher-end application of video processing, channel prices for software are in the range of USD 10,000 to USD 25,000 per stream, which again is competitive with hardware alternatives.

Density

As with cost, density varies based on the specifics of resolution, use case and selected codec. The table below provides baseline performance numbers for HPE's new m710x cartridge powered by the low-power (yet high-performance) Intel Xeon E3-1585Lv5 Processor with Iris Pro Graphics. It is worth noting that these densities reflect improvement over the baseline performance of the previous generation of processor—progress is expected to continue at an equivalent pace through future generations of the processor and the systems built around it.

Multi-stream Performance		Number of Real-time streams
1080p to 1080p, 30 fps	AVC-to-AVC	18 streams
	AVC-to-HEVC	15 streams
	HEVC-to-HEVC	8 streams
4K to 4K, 60 fps	AVC-to-HEVC	2 streams
	HEVC-to-HEVC	1 stream

In terms of simultaneous streams per rack, this translates into the ability to support 7290 simultaneous streams for transcoding 1080p30 in AVC or 810 simultaneous streams for transcoding 4Kp30 in HEVC. It should be noted that these are baseline performance numbers using the Intel MediaSDK, with all content at 8-bit 4:2:0 sampling. A vibrant community of encoder and transcoder vendors offers products that allow the adjustment of throughput and video quality to best suit individual use cases. Products are typically optimized for specific target use cases through proprietary software, which leverages Intel's acceleration features.

Power

Power consumption is competitive with hardware implementations, although not strictly speaking less than or equal to hardware average consumption. Vendors report power consumption in the range of 12.5 watts per HD stream for high-quality transcoding when hardware acceleration is fully leveraged. This is competitive with conventional implementations. For fast transcoding, power consumption as low as 3.5 watts per HD stream has been shown, which again is competitive with conventional implementations for comparable use cases; 5-6 watts per channel is considered typical for a live linear encoder of HD resolution.

Power consumption is often the most important consideration (after upfront CAPEX) in emerging markets where energy costs are high. However, it is a key factor in major markets as well. Higher power consumption leads to higher energy expenses directly and also indirectly due to increased cooling costs and less-efficient datacenter space utilization. A large studio was able to reduce data center space by 50% and power and cooling by 87% by switching to a Moonshot system for transcoding and cloud-based content transfer. Furthermore, hardware costs were reduced by 57% compared to traditional infrastructure.

Quality and Latency

In any video compression situation, three conflicting factors are quality, latency and computational horsepower. How well all three can be improved ***together*** is the measure of technological superiority of a given encoder. CPU-based systems have conventionally been considered to be at a disadvantage in terms of visual quality for a given latency and system cost. However, as GPU acceleration becomes more sophisticated and better encoding building blocks are included at silicon level, this gap is being erased. Current benchmarks show that best-in-class software systems with COTS processors can and do simultaneously deliver best-in-class performance along all three vectors—quality, latency and power consumption. This leap forward is the enabler that allows all segments of content companies to leverage virtualized systems even in highly demanding, real-time applications such as live broadcasts, Cloud DVR, and virtual reality.

Reliability

All the agility and flexibility in the world are meaningless if workflows are unreliable and live broadcasts are interrupted. Every second of an interrupted broadcast can cost millions of dollars in lost advertising revenue and adverse publicity. Traditionally, hardware was considered to be the only viable option for these highly demanding applications. Software systems today can deliver comparable reliability and can provide more efficient redundancy. For example, Moonshot-powered workflows are delivering carrier-grade reliability for single-screen and multi-screen applications for operators, OVPs and managed service providers worldwide.

Moving forward, software systems will continue to deliver higher reliability and uptime, thus offering yet another business-critical advantage over the older hardware approach. For example, the Moonshot's application-driven, hot-swappable cartridge approach increases both flexibility and reliability. As we'll discuss later, cartridges can be swapped across applications in the data center for increased redundancy without equivalent added expense. Moreover, cartridges can be swapped between the data center and the edge chassis, which further facilitates redundancy, while also controlling upfront costs.

CONSIDERATIONS IN DESIGNING A SOFTWARE-DEFINED WORKFLOW

We have seen how growth in OTT content consumption along with persistent fragmentation in rendering platforms and delivery technologies demand diversity and distribution in video processing. Designing a video workflow architecture requires choices to be made at three levels:

- Which processor is chosen on which to build a solution
- Which server or hardware appliance is chosen
- Where, physically, the processing or appliance is installed

The more flexibility one has in choosing the location of the processing, and the more efficient that processing is, the more robust and economical infrastructure becomes in the long run. "Cloud" as a generic term is often touted as a solution to needs for scalable and agile video processing. This is justified; well-designed virtualized implementations can deliver benefits such as simplified experimentation, optimized usage of infrastructure investments, and rapid spin-up/spin-down of applications and resources as required. It is important to realize, however, that "cloud" can have a variety of connotations and interpretations, and not all of these are created equal.

Data Centers versus the Public Cloud

The public cloud, typically leveraged via Infrastructure as a Service (IaaS), is a cost-efficient option for smaller operations but can prove frustrating for Tier-1 operators, large OVPs or managed service providers who need closer control over equipment, networking and distribution to achieve their requisite targets of scale and quality. Private data centers on the other hand do not have infinite room, so there is tremendous need for denser, more efficient appliances that can allow content companies to keep pace with skyrocketing growth in video traffic and video processing volumes.

Operators typically own and operate a select number of data centers complemented by more edge processing facilities. In this case, online video traffic is often delivered over the operator's own network, rather than in strict OTT fashion. Network architecture then becomes a non-trivial consideration. As discussed earlier, unicasting to individual users can quickly overwhelm networks. By moving final-stage transcoding and by processing network video functions as close to the edge as possible, the incremental cost of networking can be closely controlled and video quality can be optimized within the bounds of current network capacity. The generic cloud cannot generally be leveraged for such architectures, however.

Just-in-time transcoding (JITT) is a crucial new tool in the workflow arsenal that allows processing to be pushed farther into the edge, thus conserving network capacity while maintaining or improving end-user QoE. By trading storage capacity for increased real-time processing, JITT makes systems such as those for Cloud DVR and VOD more compact, more economical and more efficient. Processors powerful enough to support JITT sessions for thousands of users bring avant-garde services such as virtual reality and UltraHD streaming into the realm of technological and economic feasibility, helping services differentiate and delight. As content libraries become increasingly uniform across various service providers, this differentiation in quality delivered to the end user, combined with controlled cost of operations, ensures maximum profitability and growth.

Bandwidth Matters

Personalized streaming is an inherently inefficient method to disseminate popular content to millions of viewers. As more and more viewers consume on-demand content during primetime hours, neither operators nor CDNs will have enough capacity to serve them all simultaneously at the highest possible quality if streaming is implemented naively. New codecs such as HEVC, or state-of-the-art implementations of mature codecs such as AVC, can help alleviate traffic to some extent by offering improved compression efficiency. However, better compression can at most provide an improvement of a factor in the low single digits, whereas traffic projections are demanding a 10X-100X improvement in efficiency. Consequently, pushing out as much transcoding and unicast delivery to the edge as possible will be crucial in mitigating traffic problems long term. JITT in the edge, introduced in the earlier section, is a viable solution to this conundrum. However, this requires that transcoding servers be deployed in the edge and hence they must be able to withstand the demanding operating conditions of the edge. Challenges posed by this environment include hostile temperature ranges, even more spatial limitations than the head-end data center, and high reliability requirements.

Technology to the Rescue

Servers like HPE Moonshot System with the HPE ProLiant m710x server cartridge, powered by the Intel Xeon E3-1585Lv5 Processor are powerful, compact and rugged–and offer a viable solution to the demanding requirements identified above. Moonshot cartridges are interchangeable across various software stacks and between the data center chassis (HPE Moonshot 1500 chassis) and edge chassis (HPE Edgeline EL4000 Converged IoT System and EL1000). EL1000 is a ruggedized single cartridge chassis, while HPE EL4000 fits into ETSI racks in EMEA. Both chassis are designed to fit easily into outdoor broadcast trucks for live event contribution. This provides flexibility in deployment and consequently maximizes the return on investment (ROI) that customers can achieve on their CAPEX investment.

Moreover, the platform offers unique benefits such as high computational density, common sparing, high temperature tolerance, hot swappable cartridges, and low power consumption. Together, these characteristics allow the servers to be deployed reliably in the edge. Equivalently, they can be used to ramp up video processing capacity in head-ends, thus upgrading the operator's ability to handle a much higher number of channels or programs, improve resolutions and upgrade to more sophisticated compression technologies. Importantly, this expansion can be achieved within the existing spatial limitations of the head-end.

The chassis support a range of power input, including 240V, 120V and 48VDC, allowing for consistent worldwide deployment of not only the servers, but also the software running on them. The servers also meet worldwide regulatory compliance standards as well as specialized criteria such as NEBS, ETSI, PnE, and others. Thereby, these carrier-grade servers are helping usher in a new wave of possibilities and economics. As a result, they are able to support the demanding nature of best-practice architectures. Frost & Sullivan finds that a number of large managed service providers, OVPs, enterprises, and select tier-1 content providers are choosing to build rather than buy their software-based encoder/transcoder implementations. The open nature of Intel® Quick Sync Video technology for media transcoding and the high popularity of Moonshot, backed by the ability to easily customize workflows and scripts with a fully programmable, industry-standard interface (RESTful API), make Moonshot an extremely popular platform for in-house development.

Last but not least, these same factors also make Moonshot an attractive target platform for differentiated, commercial software implementations from leading video encoder and transcoder vendors. For customers who choose to buy rather than build, this rich ecosystem of vendors offers important benefits. These include flexibility to choose best-of-breed components at each point in the workflow, the assurance of competitive cost, and freedom from lock-in.

THE LAST WORD

The HPE Moonshot system featuring Intel Xeon processors revolutionized encoding and transcoding with its first release two years ago. By bringing together the performance and density once only available in hardware, with the flexibility and virtualization-friendly nature of software, the economics and feasibility of video processing for OTT were fundamentally altered. The next-generation combination of m710x with the Intel Xeon E3-1585Lv5 processor brings a new wave of progress to the table. This includes new features such as hardware-accelerated HEVC compression, support for 4K/8K, along with higher density and better throughput capacity. HPE's Edgeline EL1000 and EL4000 systems are designed for compatibility with common rack standards and fit nicely with outdoor broadcasting trucks. The combination of Moonshot cartridges and Edgeline systems significantly expands the use cases and deployment scenarios in which video and content companies can conveniently and reliably leverage Intel's game-changing processors. Frost & Sullivan research finds that Intel continues to be the processor of choice on which leading encoding/transcoding vendors are building out modern video processing technology. Frost & Sullivan best practices recommend strongly that video service providers transition to software-based workflows for optimal total cost of ownership, high agility and, ultimately, a more competitive and profitable service. Operators with software-defined workflows will be best positioned to grapple with trends such as surging OTT consumption, increasing video resolutions, continued shift toward device-centric viewing and unpredictable shifts in compression, delivery, and monetization technologies. Fortunately, technology advances from HPE and Intel are empowering operators to modernize their workflows in a performance-efficient and cost-efficient manner, thus setting the stage for growth in viewership and profitability.

FROST & SULLIVAN

NEXT STEPS

Schedule a meeting with our global team to experience our thought leadership and to integrate your ideas, opportunities and challenges into the discussion.

Interested in learning more about the topics covered in this white paper? Call us at 877.GoFrost and reference the paper you're interested in. We'll have an analyst get in touch with you.

Visit our **Digital Transformation** web page.

Attend one of our **Growth Innovation & Leadership (GIL)** events to unearth hidden growth opportunities.

SILICON VALLEY
331 E. Evelyn Ave., Suite 100
Mountain View, CA 94041
Tel 650.475.4500
Fax 650.475.1570

SAN ANTONIO
7550 West Interstate 10,
Suite 400
San Antonio, TX 78229
Tel 210.348.1000
Fax 210.348.1003

LONDON
Floor 3 - Building 5,
Chiswick Business Park,
566 Chiswick High Road,
London W4 5YF
Tel +44 (0)20 8996 8500
Fax +44 (0)20 8994 1389

877.GoFrost
myfrost@frost.com
www.frost.com