



PRESENTS

Cloud Economics: Moving TV Facilities to the Cloud in a Cost-Effective Way

Version: 1.0

Release: September 2025

Produced with the support of:

GUNPOWDER®

Executive Summary

During the past five years, the technical barriers for migrating TV and media production facilities to the cloud have been largely resolved. However, the economics of cloud migration remain opaque and unpredictable, making it difficult for executives to confidently make the transition. This whitepaper aims to shed light on the economic aspects of cloud migration, providing a detailed analysis and an accompanying cost calculator for sports, entertainment and media companies that is available for free download, at www.entertainmenttechnologist.com/cloud. The calculator was used to model various scenarios in this paper including hypotheticals that demonstrate how costs can vary by up to 50% if the workflow and technology stack is not correctly provisioned.

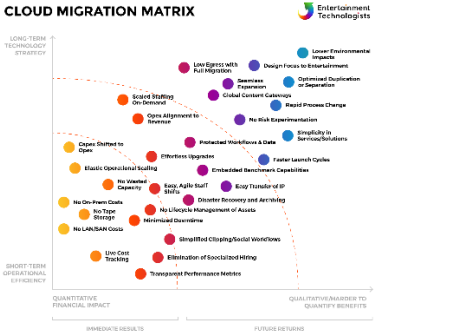
Key Points:

1. **Economic Uncertainty:** Despite transparent pricing from major cloud operators, translating their generic IT price sheets into translating media workflow ROI remains challenging. This complexity and unpredictability have led many companies to stick with traditional equipment and infrastructure, or deploy a hybrid approach which adds operational delays, balloons egress costs and adds further unpredictability. But with ongoing, and accelerating, structural change in the media and entertainment landscape, being locked into a hardware-defined workflow is too inflexible. Only those companies that are nimble and agile will succeed during this time. For that, they need a cloud-based software defined workflow.
2. **Cloud Architecture:** This paper presents a cloud architecture based on an actual broadcast facility running live news and sports networks. The architecture includes all compute, storage, and networking migrated to one cloud availability region. The design is aligned with the [10 Principles of the MovieLabs 2030 Vision](#) for modern, flexible and efficient operation of media in the cloud.
3. **Cost Comparison:** This whitepaper also compares the costs of running the facility on AWS and CoreWeave public cloud infrastructure. In the primary scenario, our sample facility would spend \$7.54 million annually on AWS or \$4.64 million annually on

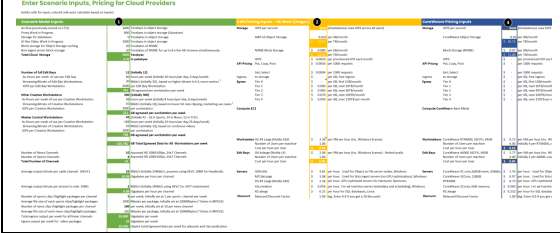
CoreWeave. CoreWeave's lower costs are primarily due to cheaper GPU-enabled machines and no data egress charges.

- 4. **Key Learnings:** This paper highlights several key learnings for cloud deployment, including the importance of accurate modeling, the critical impact of machine availability (not just pricing), the potential for cost efficiencies through GPU sharing, and other recommendations for how to optimize cloud experiences for broadcast facilities.
- 5. **Cloud Benefits:** The cloud architecture presented here offers numerous benefits, including simplified storage, flexible workflows, and the ability to quickly adapt to market changes – all summarized in an infographic in Appendix A. We have emphasized the importance of transparency and real-time cost tracking to manage cloud expenses effectively.

Conclusion: The financial and technical case for moving to the cloud can be proven both mathematically and through performance. Therefore, the decision to move media operations to the cloud is no longer about *whether to do it*, but *how soon* can it be achieved?. The economic and technical analyses provided in this whitepaper highlight the opportunities and challenges of cloud migration, emphasizing the need for careful planning and accurate modeling to achieve cost efficiencies and operational flexibility.



Appendix A – Cloud Migration Matrix of considerations when building a cloud business case.



Appendix B – Cloud Cost Calculator for Broadcast Facilities (in Excel format)

1.0 Introduction

During the last 5 years it has become clear that many of the technical barriers to the migration of TV and media production and operational facilities to the cloud have been resolved and it is relatively easy (as we have in Appendix A) to list all of the sound operational and technical reasons why any organization should move to the cloud now. However, the most vital barrier to mass cloud adoption is less frequently discussed. Simply put, **the economics of the cloud are still too opaque and too often unpredictable for many executives to stake their reputation and careers on making the migration.** One could counter and say pricing for the major cloud operators pricing is very transparent – all their pricing is posted online for all to see. However, that raw infrastructure pricing aimed at IT users is still opaque to sports, entertainment and gaming companies. Without specialist knowledge it is simply too hard to translate media workflows with complex, large, multi-location workflows into generic IT centric pricing calculators and come out with any confidence that your calculations are correct. Few technologists, let alone their finance counterparts, can calculate how many IOPS their machines use, how much data their VMs will egress once deployed or how often discs spin up, let alone what those numbers may be next year. With such pricing complexity and an unpredictable financial operating environment, is it any wonder that many companies have elected to stick with buying equipment, racking and stacking it and sweating the asset for 5 more years?

If we want to have an open and honest conversation about the economics of the cloud, and how best to optimize for it – something needs to change. So *Entertainment Technologists* leveraged our team of experienced broadcast operations managers, cloud engineers and financial analysts who have done this migration many times before, **to shine a light of honesty into the dark corners of cloud economics.**

This paper was written with technology and financial leaders at sports, entertainment and media companies in mind. We've taken a real broadcast facility running both live news and sports networks, designed a cloud architecture to migrate that facility to the cloud and calculated the costs of them doing so for two public clouds, as a comparison (AWS and CoreWeave). We're now giving the analysis and the calculator away for free. This whitepaper provides an overview of the architecture, the key technology

decisions/recommendations we've made, and shows how those decisions map directly into economic consequences by using the cost calculator to model different scenarios.

You can download the cloud calculator (in Microsoft Excel format) from our website here:

www.Entertainmenttechnologists.com/cloud and play with it yourself.

Our objective is to answer these thorny questions:

- 1) How can operators best use the cloud to power broadcast facilities and match costs to demand, in these periods of enormous change?
- 2) How can they migrate workflows and processes to the cloud, in such a way that is not disruptive to existing staff, yet allows the organization to best exploit the competitive advantages and the cloud's quickly evolving capabilities like Generative and 'Agentic' AI?
- 3) How can operators ensure long-term predictability in costs (that mimics fixed cost depreciation), when the inherent model of the cloud is to be variable in nature?

Let's dig in by looking first at the facility architecture we're looking to migrate.

2. Introduction to the Technical Use Case.

2.1 - Current System Architecture

Our target facility runs 24/7, ingesting both programming and live feeds and outputting 10+ regional news channels and 5+ sports networks. We are not going to analyze the operational costs of the facility, its staffing or software licensing costs because they are confidential, vary dramatically between regions and are known to operator (so they can compare this new model against their existing costs). We encourage operators to do the same thing; use our cost calculator (see section 3.1 for instructions) to model their own facility and see how the costs compare to what they have today. The hardest part in this process is developing an operational migration plan that accounts for ongoing running of the facility while cloud migration is occurring, especially with a need to do training of staff (change management) and onboard them on the new platforms. This is a process that requires careful planning.

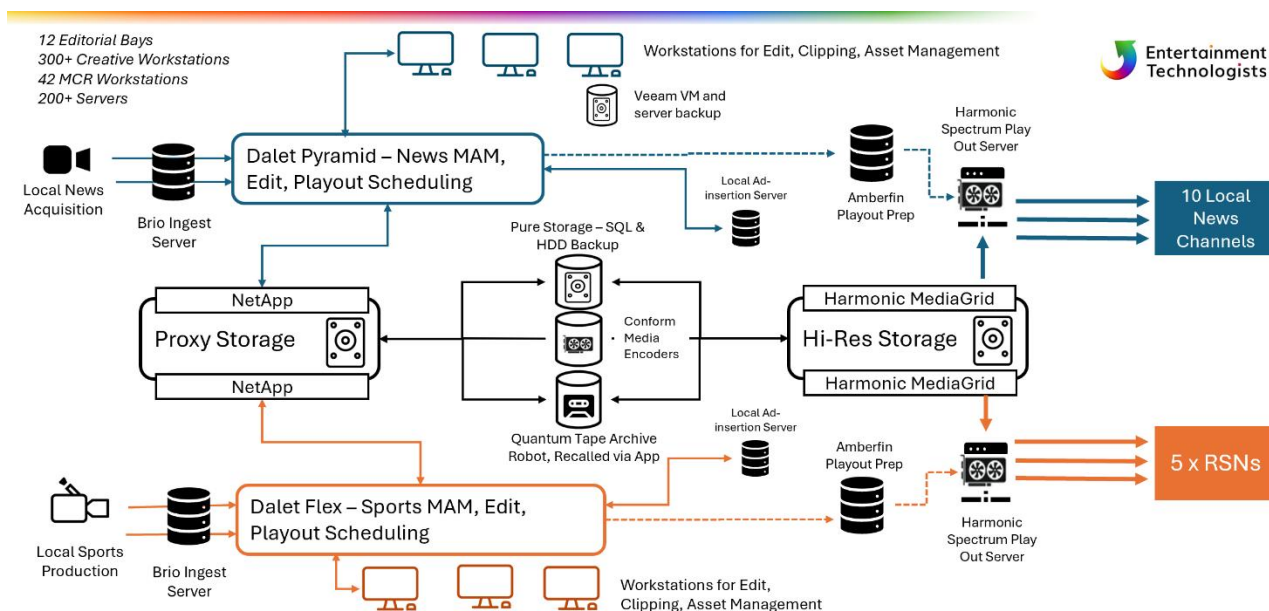


Fig 1. Existing System Architecture of our target broadcast facility that supports 15 live channels.

Our reference facility is 'medium' in size – it creates and runs regional channels in the US market which are output to both cable networks and streamed to the web. News and sports clips are syndicated out to partners and to streaming services. The facility is typical in many ways, using a Dalet MAM to manage both assets and news/sports workflows. There are in excess of 300 creative workstations used by news editors, marketing, and sports editors which are running Adobe Premier Pro to create and edit content. There are forty more 24/7 workstations accessing media from master control.

The facility requires more than 200 local servers to run the broadcast infrastructure. As you can see from the diagram – much of the costs and licensing are focused on the range of different storage solutions to support different parts of the workflow – from fast cache storage to support file ingest and conform encodes, to slower on-premises archival storage on an LTO tape robot. In total, there are around 6 petabytes of storage and considerable local networking (SMPTE 2110) to support the LAN.

All this equipment is housed on-site with redundant incoming power sources, considerable air conditioning, a 24/7 technical operation staff and regular software, hardware and routine maintenance required to support it. As they operate from a major urban center (where real estate rent is \$50-\$65/sq foot) there is considerable lost opportunity to downsize or support

more staff or stages in the ~10% of the facility square footage that is currently dedicated to technology hosting.

2.2 - Proposed Cloud Architecture

We worked with the experts at [Gunpowder](#) to re-architect this facility in order to recognize the value in migrating broadcast facilities to the cloud. While there are many options in how much infrastructure to host in the cloud, we opted here for the most extreme version – where all compute, storage, and networking are migrated to one cloud availability region. We did this primarily because most cloud benefits accrue when the data is migrated wholesale to the cloud and not moved in and out between various processes, (a “hybrid” approach that can drive up egress and increase delay). The approach we selected is consistent with the 10 Principles of the MovieLabs 2030 Vision – when all data and applications are hosted in the cloud, which enables zero-trust security and flexibility from software defined workflows. The final channels are egressed from the cloud just once to the public internet for distribution¹.

The new architecture looks like this:

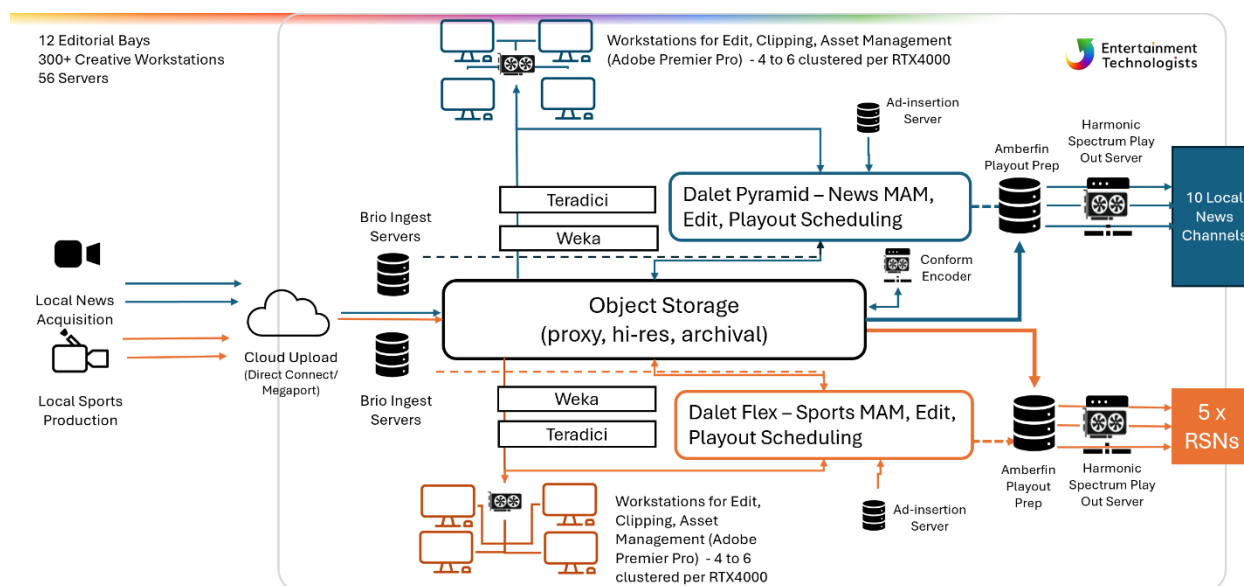


Fig 2. – Architecture for a migrated Cloud-based broadcast facility replicating the existing infrastructure with higher performance storage and compute.

¹ Our mezzanine feeds go out to cable network head ends. Although the same would be true if it was going to a streaming service or FAST channel – the video would be egressed to CDN origin storage.

2.3 - Key Technologies Introduced

Storage

Our proposed cloud architecture removes much of the legacy storage that was the 'core' of the on-premises infrastructure and replaces it with one large tier of object storage (along with removal of the costs for licensing, hardware and maintenance for all those myriad legacy storage types). This dramatically simplifies the architecture and enables an 'always-on' archive – meaning the entire story and sports archive is available to creative staff 24/7 without needing to find, request and hydrate old data from the LTO tapes^{2,3}. Also note that some functionality like growing files (from incoming live streams) are not natively supported in object storage, so are cached in our design while the feed is incoming to fast NVME storage and then written to the object storage when the event concludes. We also have an NVME layer of fast storage to accelerate the performance for the virtual workstations.

Networking

Our architecture was designed to run in one cloud data center region, but these services could also be split amongst multiple regions to provide redundancy or move data closer to end users. The physical studio facility always needed to provide dual power circuits but now also needs to provide dual and redundant high bandwidth (greater than 1 Gbit/s) internet connections to ensure cloud connectivity. However, if the studio facility lost all power and bandwidth (or was struck by a natural disaster or pandemic), in our new design it could operate entirely remotely as no local data or processing is required. If a COVID-26 pandemic occurred tomorrow, the staff could run this facility from their homes or from coffee shops anywhere in the world.

Compute & Workflows

There are other opportunities for efficiencies in this design, which we did not propose at this stage – for example, preemptible machines for certain tasks and migration of creative users to browser-based video editing (and away from Adobe Premier). We did this because many

² This speed to access data is one of those harder to measure cloud benefits, that is not detailed in our calculator but should be captured as part of a work-study by those looking to implement cloud migrations and truly reflect the operational impact and benefits.

³ The chosen MAM would need to provide proxy and searchable metadata into this object storage to allow quick find and restore facilities

of the challenges in a cloud migration are not technical, or commercial, but in *change management* with users – where it is just too disruptive to change the technologies and workflows at the same time whilst running an operating facility. That is not to say that this design is a cloud ‘lift-and-shift’– where the legacy infrastructure is just replicated in the cloud. Lift-and-Shift never results in the efficiencies that the cloud enables when properly deployed and can often drive *up* costs, diminish productivity and reduce reliability (as legacy software may not operate well in a cloud environment). Instead, we have struck a balance in this first version deployment between familiarity of workflow for the staff (easier for change management) and taking the easiest of the cloud architectures and benefits for now.

A major cloud benefit is enabling future workflow innovations, which will allow users, or even pools of users, to migrate to different technologies on a trial or permanent basis to explore future efficiencies. **Legacy broadcast facilities are typically defined by hardware-mandated workflows** – that is, the equipment routing dictates how the work must be done and changing workflows can often involve rearchitecting and rewiring racks of equipment. This approach is rigid, inflexible and locks facilities into legacy approaches, which restrict future innovation. **Modern software defined workflows, powered by the cloud, enable operators to try new software and tools (like GenAI) and open up dramatic opportunities for continuous improvement.** It is this nimbleness which is vital in the modern competitive operating environment.

Another major benefit in modern cloud deployment is the reduction in virtual machine requirements through GPU-sharing (sometimes called sharding). Our design includes virtual workstations for every user, including the master control and **GPU-sharing to dramatically lower the number of virtual machines that are required (by between 4x and 6x)** to support the entire staff. GPU-sharing uses one high powered cloud GPU instance to support multiple workers as not all users are pulling the full power of the GPU at any one time, but when they need up to 100% of the power, it is available to them. This sort of efficiency is possible with on-premises solutions but requires buying hardware to support maximum capacity, which then becomes a capex burden and a potential cap on future hiring. The modern cloud approach allows GPU instances to be expanded during peak times (e.g., major events) when staff scales up and immediately shut down when the work is done, or during overnight shifts. The matching of costs to demand is one of the key objectives in

our design and can support facilities hiring new workers or releasing workers, while ensuring costs are dynamically tracked.

Integration with existing tools – Another key to user stability in our design is that the Dalet MAM and workflow management systems, [Flex](#) & [Pyramid](#), run natively in the cloud. This enables staff to continue using familiar tools and workflows while the underlying infrastructure is dramatically reengineered. By connecting the object storage directly to Dalet, we can also remove multiple layers of software, costs⁴, and staffing to manage that storage, and speed up performance. There are additional software licensing costs for this new design which partially offset the removal of the local storage management, including costs for running remote workstation streaming software (we have proposed HP Anywhere, formerly Teradici) and a fast cache block storage system to accelerate the performance of object storage. There are multiple providers of these file acceleration services, we have designed around [Weka](#), but their operations are similar – they provide file gateways to bridge legacy applications into object storage and provide a fast cache to accelerate performance. In our model, we have specified 20 terabytes of high speed NVME storage to cache and accelerate performance of the 6 petabytes of object storage. These two software layers accounted for a large proportion of the costs of the cloud migration, but our calculator uses ratecard pricing. At the volumes we are proposing, all vendors will provide discounts for scale. So, these costs should be viewed in our calculator as a worse-case scenarios.

2.3 Cloud Architecture Benefits

Rather than list and detail out all of the benefits and considerations of migrating broadcast facilities to the cloud, we have condensed them down into this friendly infographic:

⁴ Cutting ~100 servers and associated licenses out of the design

CLOUD MIGRATION MATRIX

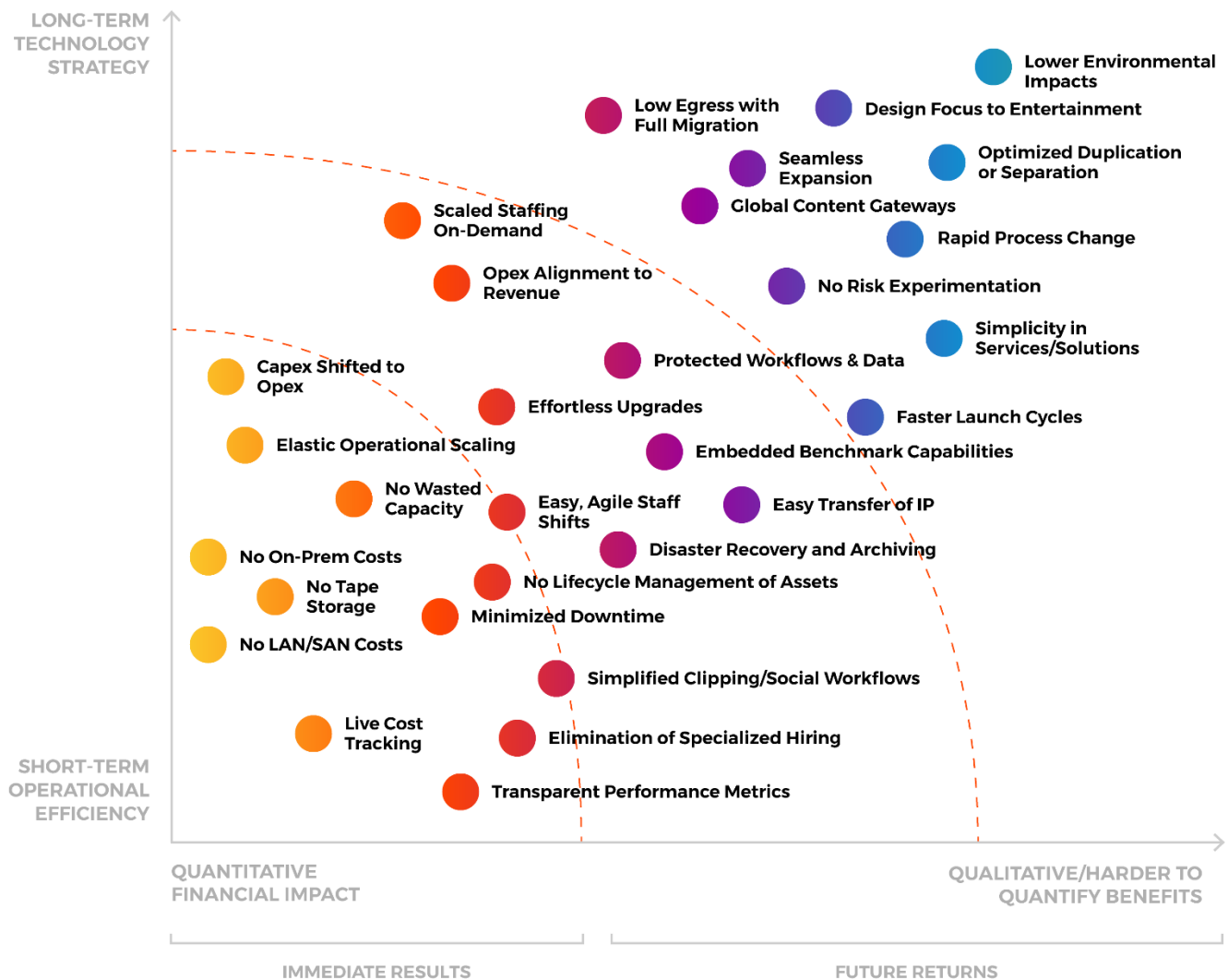


Fig 2. Cloud Migration Matrix plotting the considerations (fiscal, technical, operational) for those considering a migration.

See more details in Appendix A.

Immediate, tangible, near-term benefits are represented in the lower left section of the infographic. While other cloud benefits are harder to immediately measure (e.g., future nimbleness to adapt to market change) or even apply a specific value to (e.g., quickly recovering from a major disaster). Those values more contingent on management decisions or future events and are represented further out towards the top-right in our infographic. That does not mean they have less importance – some may contain the greatest overall benefits to an organization. It is just a matter of how and when their value will be realized and what competitive posture an organization might take after deploying to the

cloud. **When finance, technology and operations teams meet to discuss a cloud migration, this diagram lists the areas that should be discussed and mapped out, according to where each individual organization ranks their particular importance.** As we are focused here on near-term financial cost modelling, we won't dwell on these considerations but are happy to run workshops to discuss in more detail how they can be assigned value for each organization and how to ensure the organization is set up to fully capture these.

The full version of the matrix is included in Appendix A.

3. The Cloud Cost Calculator

This whitepaper is accompanied by a full Excel cost calculator that Entertainment Technologists is providing for free for the community to download, use and modify for their own scenarios – download at www.entertainmenttechnologists.com/cloud. We have taken our reference facility and cloud architecture and calculated the costs of running that facility with two different vendors – AWS (in US-West, Oregon) and for comparison a newer cloud operator CoreWeave (in their Las Vegas data center).

3.1 – How to Use the Cloud Cost Calculator

Enter Scenario Inputs, Pricing for Cloud Providers

(white cells for input, colored cells auto-calculate based on inputs)

Scenario Model Inputs		AWS Pricing Inputs - US West (Oregon)				CoreWeave Pricing Inputs			
Active (previously stored on LTO) Proxy Work in Progress Storage for databases Hi-Res Video Work in Progress Block storage for Object Storage caching Brio ingest server block storage Total Cloud Storage	3000	Storage IOPS per second		6600	(simultaneous max IOPS across all users)		Storage IOPS per second	6600	(simultaneous max IOPS)
	8000								
	0	AWS S3 Object Storage		\$ 0.023	per GB/month		CoreWeave Object Storage	\$ 0.03	per GB/month
	20000			\$ - 215.5	per TB/month			\$ 30.72	per TB/month
	20								
	200	NVME Block Storage		\$ 0.080	per GB/month		Block Storage (NVME)	\$ 0.07	per GB/month
	200			\$ 0.0050	per TB/month				
	200				per provisioned IOPS each month		IOPS		per provisioned IOPS each
	5880	API Pricing Put, Copy, Post		\$ 0.0050	per 1000 requests		API Pricing Put, Copy, Post		per 1000 requests
	5.20								
Number of full Edit Bays Av Hours per week of use per Edit bay Streaming Rate of Edit Bay Workstation IOPS per Edit Bay Workstation Other Creative Workstations Av Hours per week of use per Creative Workstation Streaming Rate of Creative Workstations IOPS per Creative Workstation Master Control Workstations Av Hours per week of use per Creative Workstation Streaming Rate of Creative Workstations IOPS per Creative Workstation	12	Get, Select		\$ 0.0004	per 1000 requests		Get, Select		per 1000 requests
	60	to storage		\$ -	per GB, free ingress		Ingress		per GB, free ingress
	30	Egress Tier 0		\$ -	per GB, first 1GB/month		Egress Tier 0		per GB, first 1GB/month
	4000	per Edit Bay Workstation		\$ 0.090	per GB, next 10TB/month				per GB, next 10TB/month
	780	GB egressed per workstation per week		\$ 0.085	per GB, next 40TB/month				per GB, next 40TB/month
	300	Tier 3		\$ 0.070	per GB, next 100TB/month				per GB, next 100TB/month
	40	hours per week (initially 8 hours/day, 6 days/week)		\$ 0.050	per GB, over 150TB per month				per GB, over 150TB per month
	10	MiB/s (initially 10), based on lower bitrate 4:4:4, more motion ¹							
	10	MiB/s (initially 10), based on lower bitrate 4:4:4, more motion ¹							
	2000	per workstation					Compute EC2		Compute CoreWeave Bare Metal
Number of News Channels Total Number of Channels Average output bitrate per cable channel (HEVC) Average output bitrate per stream to web (VBR) Number of sports clips/highlight packages per channel Average file size of each sports clip/highlight package Number of news clips/highlight packages per channel Average file size of each news clip/highlight package Total gross output per week for all linear channels Egress output per week for video packages	40	GB egressed per workstation per week		\$ 0.085	per 1000 requests				
	168	Initials: 42, 16 in Sports, 14 in News, 12 in TOC)							
	168	hours per week (initially 24 hours/per day, 76 days/week)							
	10	MiB/s (initially 10), based on continuous videos							
	2000	per workstation							
	780	GB egressed per workstation per week							
	103,761	GR Total Egressed Data for All Workstations per week							
	10	Assumed HD 1080i 60fps, 24x7 Channels					Workstations		Workstations
	5	Assumed HD 1080i 60fps, 24x7 Channels							
	15	Cost per hour per User		\$ 2.36	per VM per hour (inc. Windows license)				
Average output bitrate per cable channel (HEVC) Average output bitrate per stream to web (VBR) Number of sports clips/highlight packages per channel Average file size of each sports clip/highlight package Number of news clips/highlight packages per channel Average file size of each news clip/highlight package Total gross output per week for all linear channels Egress output per week for video packages	25	Number of Users per machine		1.00	Number of Users per machine		Number of Users per machine	4.00	Initially 4 per RTX4000,...
	0	Cost per hour per User		\$ 2.56	per VM per hour (inc. Windows license) - limited avails		Cost per hour per User	\$ 0.18	per VM per hour (inc. Windows license) - limited avails
	10	Assumed HD 1080i 60fps, 24x7 Channels		\$ 2.06	per VM per hour (inc. Windows license) - limited avails		Cost per hour per User	\$ 1.77	per VM per hour (inc. Windows license) - limited avails
	5	Assumed HD 1080i 60fps, 24x7 Channels		1.00	Number of Users per machine		Number of Users per machine	2.00	Initially 2 per A6000, cost...
	15	Cost per hour per User		\$ 2.56	per VM per hour (inc. Windows license) - limited avails		Cost per hour per User	\$ 0.88	per VM per hour (inc. Windows license) - limited avails
	25	Number of Users per machine		\$ 3.81	per hour. Used for Object to File server nodes, Windows		Servers		Servers
	0	Cost per hour per User		\$ 3.08	per hour. Used for Brio ingest servers (no GPU optimization), Windows				
	10	Assumed HD 1080i 60fps, 24x7 Channels		\$ 2.36	per hour. GPU optimized servers for Harmonic Spectrum, Windows		CoreWeave 41c, core156GB mem, 2GBdss	\$ 1.43	per hour. Used for Brio ingest...
	0	Cost per hour per User		\$ 0.056	per hour. For ad-insertion servers (metastad and scheduling), Windows		CoreWeave 32Core, 128GB	\$ 0.97	per hour. Used for Object to...
	10	Assumed HD 1080i 60fps, 24x7 Channels		\$ 0.25	per hour. For ad-insertion servers (metastad and scheduling), Windows		RTX4000	\$ 0.73	per hour. GPU optimized...

Fig 3. Inputs tab of the free cloud calculator for broadcast facilities.

Our calculator is easy to use and includes instructions and key assumptions on the first tab. It is important to understand the assumptions because they can have a large impact on the

model and if they do not match your scenario, the calculator should be modified. For example, we've calculated egress to the public internet (which, allows our cloud facility to operate from anywhere in the world), but the AWS costs should be changed if you are connecting a bricks-and-mortar facility with a direct-connect connection to AWS as that will lower egress costs.

On the "Input" tab, users can enter data into the white spaces (the colored cells auto-calculate). We include baseline values so users can always get back to the base case if they have been running alternate scenarios. There are 3 sections on this tab:

- The first section in **Green** includes all the key assumptions about the target facilities(s) including the amount of storage, number of user workstations, servers, channels, playout bitrates etc.
- The **Gold** section includes the pricing for first cloud provider, in our case AWS. We have used ratecard pricing for US-West as our baseline, but you can change to pricing from a different region by modifying these cells. If you have a bulk discount with AWS (eg. by making a spending commitment), enter that in cell H4I as a discount factor. (i.e., entering a value of 0.9 here reduces pricing by 10%). The pricing is calculated both with and without the discount so you can see the impact it has.
- The **Blue** section includes pricing for our 2nd cloud provider, CoreWeave. They do not charge different pricing per region or for data egress. So, some of these cells are set at zero. There is still the ability to apply a discount factor by modifying cell M4I.

The rest of the calculator runs automatically. Each cloud operator has its own tab showing detailed calculations for the relevant scenario and pricing. The final tab provides a summary of the costs of the relevant scenario for each cloud provider and compares them against each other. It is possible to add additional cloud operators for comparison, but this must be done carefully as each provider has unique pricing structures and rules which impacts storage tiers, machine types, performance characteristics and other factors. The baseline cloud calculator we have provided is accurate for a like-to-like comparison of these two providers but may differ if others need to be compared. We recommend deeper scenarios or ones involving additional providers be formulated in conjunction with *Entertainment Technologists*.

3.2 – Comparison of 2 Public Cloud Vendors

AWS started the public cloud marketplace and is still the dominant provider with a myriad of services and vendors in their AWS Marketplace and a global footprint of data centers (Availability Zones). AWS has developed reference architectures for many sports, entertainment, and gaming use cases and almost all applications and SaaS services are designed to operate on AWS. Amazon has a customer centric model and frequently promotes they regularly and proactively lower their prices. However, it is worth noting that since we last did our cloud cost calculations in 2020, AWS has dropped their pricing in only one key category as shown below:

S3 Object Storage	No Change
File Storage (EBS)	16% reduction
Data Egress to the Internet	No Change
EC2 Instance (g4dn.4xlarge)	No Change

Given that over the past 5 years, the price of hard disc drives has dropped by 30–50%, and NVME storage has dropped by a whopping 60–70%, it does not appear that AWS is passing on these cost savings to customers. Similarly, compute performance in the last 5 years has dramatically improved. GPU performance now far exceeds Moore’s law. Yet, it does not seem that all of that value is being passed on to customers. Of course, raw hardware costs are not the only cost in operating a cloud⁵ but, it does seem that AWS claims for driving up customer value and lowering costs are, at best, overstated.

CoreWeave meanwhile is a relatively new provider whose data centers have largely been built over just the past 2 years. CoreWeave positions itself as the ‘AI Cloud’ and they certainly have access to a huge number of NVIDIA GPUs⁶ and deploy them faster than most other operators. Although seen as primarily a GPU cloud, we also found that CoreWeave has all the services of other public clouds including object storage, high performance disc storage, modern CPU instances and fast networking between regions. CoreWeave data centers are focused in North America but they are now expanding to Europe as well. CoreWeave also

⁵ Although networking costs and data center costs have also been dropping. Plus many of these large data centers now run on renewable energy which reduces costs still further.

⁶ NVIDIA is a large shareholder of CoreWeave and seem to guarantee availability of the latest GPUs to CoreWeave as well as reference designs for efficient data center operations.

makes modelling costs much easier and predictable because they don't charge for egress of data, nor API or other data access costs, and the pricing is the same in all cloud regions.

3.3 – Key Takeaways and Scenario Testing

By using our reference facility, architecture and financial model we can do some comparisons between cloud vendors and also test some scenarios to extract key economic learnings.

The initial results of our analysis, with like-for-like comparisons of machine types and storage, shows:

OUTPUT COMPARISONS – ANNUAL

	AWS		CoreWeave		Difference	Delta (%)
Storage Costs	\$	1,678,541	\$	2,172,518	493,978	29%
Compute Costs						
Workstations	\$	2,710,177	\$	880,739	(1,829,438)	-68%
Servers	\$	1,187,055	\$	381,938	(805,117)	-68%
Egress	\$	116,497	\$	-	(116,497)	-100%
Licensing & Other	\$	1,845,288	\$	1,845,288	0	0%
Annual without Discount	\$	7,537,558	\$	5,280,483	(2,257,075)	-30%
Blanket discount amount		0%		0%		
Annual with Discount	\$	7,537,558	\$	5,280,483	(2,257,075)	-30%
Per Month with Discount	\$	628,130	\$	440,040	(188,090)	-30%

Table 1 – Like-for-like comparison of AWS and CoreWeave costs at rate card pricing

This model shows the cost of moving to the cloud to be \$7.54m annually on AWS and \$5.28m annually on CoreWeave. The discrepancy in pricing is interesting with CoreWeave costing more on storage (29% more than the AWS ratecard) but overall coming out 30% less expensive because of the lack of egress and far reduced costs for GPU-enabled machines (68% less). As the cost of compute is dominant in our model with ~350 workstations running, the compute cost savings add up.

In both models, one of the larger costs is the ratecard charge for object-to-file management clusters (\$1.85m). These costs represent a high percentage of the overall costs on the CoreWeave model, (35%) but are lower than the on-premises licensing costs

of various storage management solutions needed to run the legacy infrastructure. As noted earlier, the costs of these cloud storage management solutions can be negotiated down as there are multiple vendors. In addition, this scenario would be considered a large deployment which, would be unlikely to be priced at standard ratecard levels.

Here are 4 other key learnings you can take from this model by adjusting the parameters and trying different scenarios:

1) Unpredictability - Justifiably the Bane of Finance Teams

One of the criticisms of public cloud is that its costs are unpredictable and that often the first time a finance department sees that a project has gone over budget is when they receive the invoice, at which point it's too late to correct.

For example, our base model assumes that creative workstations are running for an average of 48 hours per week (6 days at 8hours/day) and the edit bays at 60 hours (10 hours per day for 6 days)⁷ and that when the users go home, they switch off the machines or we detect no usage and remotely shut them down. But if you don't implement that proactive management of machines and let them run 24/7 (by changing D15 and D20 to 168), we see a dramatically different result:

OUTPUT COMPARISONS – Workstations running 24/7

	AWS		CoreWeave		Difference	Delta (%)
Storage Costs	\$	1,678,541	\$	2,172,518	493,978	29%
Compute Costs						
Workstations	\$	7,266,870	\$	2,366,582	(4,900,288)	-67%
Servers	\$	1,187,055	\$	381,938	(805,117)	-68%
Egress	\$	123,600	\$	-	(123,600)	-100%
Licensing & Other	\$	1,845,288	\$	1,845,288	0	0%
Annual without Discount	\$	12,101,353	\$	6,766,327	(5,335,027)	-44%
Blanket discount amount		0%		0%		
Annual with Discount	\$	12,101,353	\$	6,766,327	(5,335,027)	-44%
Per Month with Discount	\$	1,008,446	\$	563,861	(444,586)	-44%

⁷ The 40 workstations in MCR are running 24/7 in our model.

AWS costs, which are heavily impacted by GPU costs, rise by a budget busting \$4.5m – a 37% increase in costs that no finance department will want to hear. CoreWeave costs rose \$1.5m (22%) because their GPU costs are so much lower at the outset. While an extreme scenario, it does show **the care that must be taken to ensure that the variable nature of cloud costs work for you, and not against you.**

Cloud providers do provide near real-time cost tracking APIs and there are various companies that provide tools that utilize those to alert customers when their costs are accelerating, increasing rapidly or exceeding set budgets. We highly recommend utilizing such services and also giving users agency in their projects by exposing the costs of running machines, servers and projects in real-time. An important part of change management is to educate users that they are running in the cloud and not on a local machine. All staff, end users and engineers alike, need to relearn how they use software and solutions to take advantage of the cloud migration and one way to do that is to provide desktop level tools and data that shows the impact of the work they are doing on the operational costs of the business. **We find transparency leads to empowerment of the team and better results.**

2. More Powerful Machines, Don't Always Cost More

Although our base model includes an apples vs apples cost comparison of AWS versus CoreWeave, we actually don't recommend this architecture. We specified AWS GPU enabled machines (running Nvidia A10 GPUs) for creative workstations that are acceptable for Adobe Premier Pro users, however they are far older and slower than the base machines we selected for CoreWeave (the newer and more powerful Nvidia RTX4000 GPUs). In fact, **the CoreWeave GPU machines can each easily support 4 creative users on Premier Pro and possibly up to 6 users** (based on whether they are working on 1080, 4K or high bitrate video). The Nvidia RTX4000 GPUs CoreWeave provides are cheaper than the alternative provided by AWS and this cost difference is multiplied when they are deployed for shared use. This difference can be calculated by changing cell M30 in our cost calculator from 1 to 4 or, upto 6.. With six users per machine, the CoreWeave costs for compute drop to just 18 cents per user per hour. In that scenario you can see the total costs for CoreWeave workstations drops to just \$236,000 per year across all users – less than 10% of the costs of workstations on AWS.

OUTPUT COMPARISONS – GPU Sharing on modern GPUs

	AWS	CoreWeave	Difference	Delta (%)
Storage Costs	\$ 1,678,541	\$ 2,172,518	493,978	29%
Compute Costs				
Workstations	\$ 2,710,177	\$ 236,752	(2,473,425)	-91%
Servers	\$ 1,187,055	\$ 381,938	(805,117)	-68%
Egress	\$ 116,497	\$ -	(116,497)	-100%
Licensing & Other	\$ 1,845,288	\$ 1,845,288	0	0%
Annual without Discount	\$ 7,537,558	\$ 4,636,496	(2,901,061)	-38%
Blanket discount amount	0%	0%		
Annual with Discount	\$ 7,537,558	\$ 4,636,496	(2,901,061)	-38%
Per Month with Discount	\$ 628,130	\$ 386,375	(241,755)	-38%

Operators will want to do their own testing with creative users during both off-peak and peak-load hours and seasons to establish how many users they can run on each GPU machine. This is one of the areas of innovation that can be best utilized in the cloud and certainly opens up huge cost efficiencies if engineers can break the linearity between 1 machine for 1 user.

3. Pricing is Important, but Availability May be Key

Some readers may object to the comparison above and point out that AWS does have more modern GPUs in their fleet. For example, one might point to the G6 instances instead of the G5 instances we specified for creative workstations. We did specify G6 machines for our editorial workstations and they are actually cheaper than G5s as well as being more powerful. However, there is a key reason we didn't recommend them for the 300 creative workstations. We only needed twelve of the G6 machines to power the editorial workstations in our scenario, but even that would be a considerable stretch in some AWS regions of the US and the world. At the time of writing (September 2025) G6 instances are simply not available in sufficient quantities across AWS regions and this scarcity is expected to continue for an unknown but significant amount of time. Since the rise of GenAI and the massive new demand for GPUs for AI training and inference, it has become a frequent and persistent problem that major cloud providers simply do not have sufficient capacity for the Sports, Entertainment and Gaming industries. This leads us to our next major takeaway – **we**

were told that public cloud vendors would effectively provide limitless computing, and that is unfortunately not true.

We worked on cloud burst rendering for major animated and VFX heavy movies and when a major project was finishing cloud providers would run out of machines across whole regions, especially those with preemptible prices, as the render would simply absorb 10,000's of cores swamping the cloud region. But those events were largely predictable and happened fairly infrequently. The more persistent problem now is an ongoing shortage of GPUs across clouds and regions for use in creative industries.

In some cloud architectures we recommend pricing in preemptible machines (the "spot fleet" as they are sometimes called) where cloud companies sell off excess capacity at lower prices (30% or more) but on the understanding they can evict a tenant and reclaim the capacity if they need it. We did not recommend it in the broadcast use case here because it's just not possible to have a master control room preempted because a higher paying customer comes along. However, there are some use cases where it could be more cost efficient to use preemptible machines (like running overnight batch jobs of video encodes or composites) but in the current climate we're not recommending it because major GPU training runs at a silicon valley startup can suddenly remove all available capacity in the spot fleet causing preemption of the broadcast facility.

4. Accurate Modelling Is Key

There are numerous other scenarios you can run using the calculator to see what the impact could be. In cases where Operations has worked closely with Finance to develop a cloud business case, based on specific assumptions, outliers often occur and result in very uncomfortable discussion with finance at quarters end. The key to avoiding nasty financial surprises is to accurately model the facility, the workflow, and the working style of the users *before* doing a cloud migration. For example, below are some simple changes that have big net results in our calculator. All of these are real examples from cloud migrations we have experienced that could have been avoided if all of these areas that drive costs were accurately and thoroughly investigated and modeled:

- Doubling the broadcast master stream bitrate to 50Mbit/s and the number of channels, increases AWS egress costs by over \$250,000 per year.
- Moving the high-bit rate proxy files (about 2 petabytes) from object storage to higher performance NVME storage increases the total cost of the scenario to \$9.8m for AWS

and \$5.8m for CoreWeave but also removes the delta between their two storage costs. Both AWS and CoreWeave in that scenario cost around \$3.2m in storage annually – the discrepancy is because CoreWeave charges more than AWS for object storage but less for faster NVME storage.

- Active archives (where the entire archive can be egressed every 3 months) can drive costs up dramatically. If this facility egressed their archive every 3 months that would be around 1PB per month and the cost for AWS egress suddenly rises to \$720k per year – a \$600,000 increase. Accordingly, any big data migration, or a change in rights which facilitates a big movement of data becomes a huge egress expense. Companies that sell or syndicate their content can license hundreds of high-resolution master titles or thousands of clips per year and they should model very accurately how their cloud provider charges for egress.

But on the bright side, the existence of a model does allow the Finance, Operations and Technology teams to have honest and transparent discussions about different scenarios and which technologies will have the biggest ROI, both now and into the future.

Conclusion and Next Steps

There is as much variability in cloud design and potential optimizations as there ever was in on-premises hardware. The challenge is not that cloud architecture is necessarily harder but it's just very *different*. There are many reasons (see Appendix A) to move operations of broadcast facilities to the cloud and certainly now economic reasons seem to outweigh the inertia of doing things "like we always did". Now **the decision becomes not should we move to the cloud, but how soon it can be achieved?**

We provide this economic and technical analysis to highlight some of the technical and financial opportunities to prepare for a world which will only continue to accelerate in its change. Only facilities that are nimble will survive and even thrive.

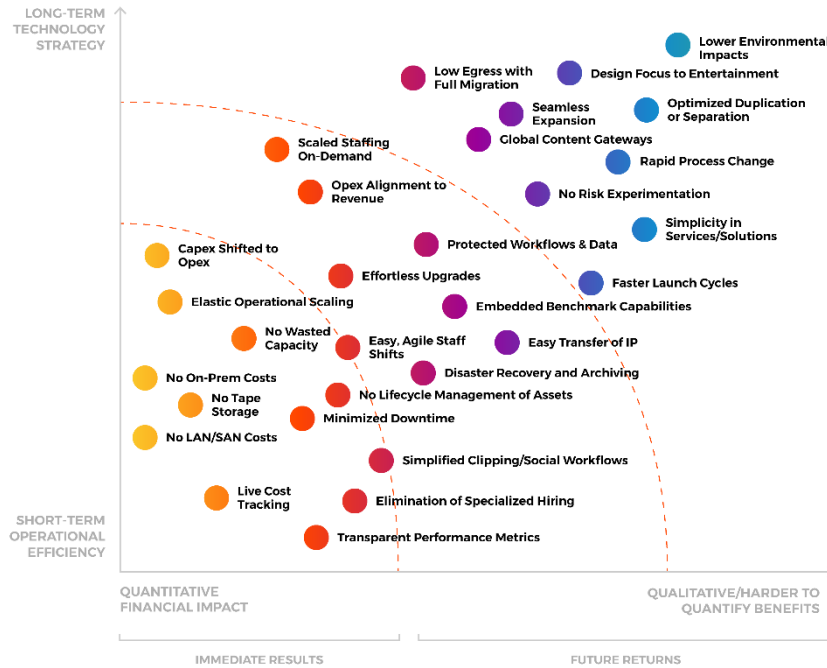
It is also clear now (which was not always the case) that there are considerable cost differences between major cloud providers and that new entrants have forced costs down through competition, to the point that it is now very economically viable to move entire facilities and networks to the cloud while improving their capabilities. Deciding on a cloud partner should be based not just on their pricing but also the technical support, their partner

ecosystem, and the assistance they are able to provide in planning and executing that complex migration.

We welcome you to develop your own use case based on our calculator or contact *Entertainment Technologists* to help build your own design, workflows and economic models.

Appendix A

CLOUD MIGRATION MATRIX



KEY

No On-Prem Costs	On-premises equipment requires floor space (opportunity cost), expensive electricity (as often is cited), and requires cooling - HVAC are big bulky and require frequent maintenance.
No LAN/SAN Costs	Proprietary labor, wiring, routers, software, maintenance, etc. are eliminated.
Capex Shifted to Opex	On-prem data centers require large, difficult to manage capex that shifts to opex in the cloud.
Elastic Operational Scaling	Variable workloads can spin up or down with ease and unit economics improve for bursty demand.
No Tape Storage	On-prem tape storage requires large amounts of space - upgrades to tape stock - every 3 years.
Live Cost Tracking	All operational costs in the cloud are fully transparent and updated through flexible dashboards.
No Wasted Capacity	Unlike on-prem data centers, compute and infrastructure grow as needed.
Scaled Staffing On-Demand	Staff flexing is now possible - match the staffing levels to the cloud and they can burst together by bringing augmenting staff for major events, busy periods.
Opex Alignment to Revenue	As the organization grows or contracts, costs move in sync to evolving needs.
Transparent Performance Metrics	Performance metrics in the cloud are constantly monitored and configurable.
Minimized Downtime	Downtime and risk from infrastructure/software upgrades is minimized and less expensive than on-prem solutions.
Easy, Agile Staff Shifts	Staff can be shifted to higher priority areas without reassigning jobs or reorganizing projects.
No Lifecycle Management of Assets	No need to pre-budget for hardware refreshes or manage assets beyond their usable life hoping they don't fail this year.
Effortless Upgrades	Upgrades to servers and software are applied automatically without action from the operator.
Elimination of Specialized Hiring	Cloud deployments do not need to hire engineers 24/7 when expert staff may not be available.
Simplified Clipping/Social Workflows	Content is more easily clipped and posted socially using cloud operations.
Low Egress with Full Migration	Media for cloud distribution should be made in the cloud. With full integration, data and content do not egress until publication of content.
Disaster Recovery and Archiving	Disaster recovery and continuity - Selfly literally just continues to operate as long as you can get an internet connection. Protection from fire, flood, power outage, pandemic, etc.
Protected Workflows & Data	Cloud operations are automatically secure and backed up.
Embedded Benchmark Capabilities	Cloud provides access to emerging technologies like GenAI, AI indexing, and other modern tools not available in legacy on-prem infrastructure.
Easy Transfer of IP	Newly licensed, expired or transferred content (and/or data) is easily moved in or out of the system.
Global Content Gateways	Content producers or providers can deliver content to the system easily from any location worldwide.
Seamless Expansion	As products and staff shift to new regions, workflows and infrastructure can seamlessly shift with it.
No Risk Experimentation	Cloud and Software Defined Workflows allow nimbleness and ability to try and rapidly deploy new ideas and experiment with new workflows and approaches.
Design Focus to Entertainment	Various cloud infrastructure elements and workflows built specifically for media operations.
Faster Launch Cycles	Technology is no longer the constraining factor in adopting new business models, growing into new markets or developing new offerings.
Simplicity in Services/Solutions	As optimal configurations are developed, alternative ones can copy or merge to those models using readily available cloud solutions.
Rapid Process Change	Modifications can be made to workflows and implemented with minimal disruptions.
Optimized Duplication or Separation	As organizations merge or split, operations and content can be easily shifted.
Lower Environmental Impacts	Unlike cloud, on-prem equipment requires expensive electricity and largely dirty energy.

Appendix B – Cloud Cost Calculator

Available for download at www.entertainmenttechnologists.com/cloud

Enter Scenario Inputs, Pricing for Cloud Providers

(white cells for input, colored cells auto-calculate based on inputs)

Scenario Model Inputs		AWS Pricing Inputs - US West (Oregon)		CoreWeave Pricing Inputs	
1		2		3	
Archive (previously stored on LTO)		Storage IOPS per second		Storage IOPS per second	
Proxy Work in Progress		Simultaneous max IOPS across all users		Simultaneous max IOPS	
Storage for databases		AWS S3 Object Storage		CoreWeave Object Storage	
Hi-Res Video Work in Progress		NVM Block Storage		Block Storage (NVME)	
Block storage for Object Storage caching		API Pricing Put, Copy, Post		API Pricing Put, Copy, Post	
Brio Ingest server block storage		Ingress to storage		Ingress to storage	
Total Cloud Storage		Egress Tier 0		Egress Tier 0	
Number of full Edit Bays		Tier 1		Tier 1	
As Hours per week of use per Edit Bay		Tier 2		Tier 2	
Streaming Bitrate of Edit Bay Workstations		Tier 3		Tier 3	
IOPS per Edit Bay Workstation		Tier 4		Tier 4	
Other Creative Workstations		Compute EC2		Compute CoreWeave Bare Metal	
As Hours per week of use per Creative Workstation		Workstation (5.4K Large (Nvidia A10)		Workstations CoreWeave RTX4000, 16CPU, 64GB	
Streaming Bitrate of Creative Workstations		Edit Bays		Edit Bays CoreWeave A6000 16CPU, 64GB	
IOPS per Creative Workstation		Servers		Servers CoreWeave 41 cores, 160GB mem, 20GBdis	
Master Control Workstations		Discount		Discount	
As Hours per week of use per Creative Workstation					
Streaming Bitrate of Creative Workstations					
IOPS per Creative Workstation					
Number of News Channels					
Number of Sports Channels					
Total Number of Channels					
Average output bitrate per cable channel (HEVC)					
Average output bitrate per stream to web (H.264)					
Number of sports clips /highlight packages per channel					
Average file size of each sports clips /highlight packages					
Number of news clips /highlight packages per channel					
Average file size of each news clips /highlight packages					
Total egress output per week for all linear channels					
Egress output per week for video packages					