THE BUSINESS BENEFITS OF OBJECT STORAGE

AND WHY IT WILL REPLACE FILE STORAGE

CONCURRENT
a vecima company

TABLE OF CONTENTS

OBJECT STORAGE

Object storage. A relatively new storage technology that promises big benefits for Media businesses. You may have heard of it and be curious about what it can do for you. This e-book explains the key differences between object-based storage vs. file-based storage and how these differences can affect your business results.

Object storage is an alternative to file-based disk storage systems and also to tape-based storage systems. It has emerged during the current Information Era to support large and rapidly-growing content libraries. It has specifically been targeted at "unstructured data", such as documents, images and videos, as opposed to "structured data" which include databases for financial or other numerical data.

In the Media industry, the incumbent storage platforms are tape-based and file-based, and they have evolved over many years of analog and digital content production and distribution. As the industry looks to the future, there are business-critical decisions to make about how content is stored. Not just to keep it secure, but also to make it accessible.

There are technical differences between these 3 alternative storage systems, which manifest in different operational capabilities. As Media businesses look to the future, operational capabilities need to be well positioned to cater for the growth in content and the growth in distribution outlets. The right choice of storage platform will contribute to your basic competitiveness and ability to leverage your content for best results.

PLANNING FOR SUCCESS

The nature of your next-generation storage is a business-critical decision for Media businesses for various reasons. Before thinking about storage, it is first vital to consider business targets and the necessary operational capabilities. In the Media industry, global trends are driving decisions about financial investments, operational priorities and technology choices, including:

CONTENT MANAGEMENT

- **Content libraries are growing continuously**
  Normal content production and archiving practices add more content to content libraries continuously.
- **Content growth rate is accelerating**
  Larger content formats such as 4K and VR make each file larger than their HD and SD equivalents.
- **Content migrations need to become extinct**
  There is a high perennial drain on an organization to manage content migrations from one file or tape-based storage format to the next, and it will only get worse as libraries grow.
- **Content libraries must perform as they grow**
  As storage systems grow in size, they need to maintain their throughput performance to support operational demands. Legacy storage systems have architectural limitations which create throughput bottlenecks, generally through file read/write devices and controllers, and robotic equipment.

CONTENT MONETIZATION

- **Content must be accessible to be monetized**
  Making money from content is the business. Storage locations and media management systems that lock down content prevent monetization.
- **On-demand actually means "now"**
  Consumer expectations for accessing content have become higher and higher over the last 5-10 years. If content is "available" in a search, but actually not available immediately, the customer experience is poor. In addition, Media Operations needing to access content immediately can't afford to wait for physical resources to catch up.
- **Metadata for content search is business critical**
  In a crowded world of content, search-ability is a key business differentiator. Creating useful metadata about your content and storing as much as you desire to make your content easily searchable could differentiate you from the competition. Metadata extensibility and protection are business must-haves.

Against the backdrop of these current trends, every Media business targeting excellent business performance needs a Media Operation that can deliver 3 key outcomes in the future:

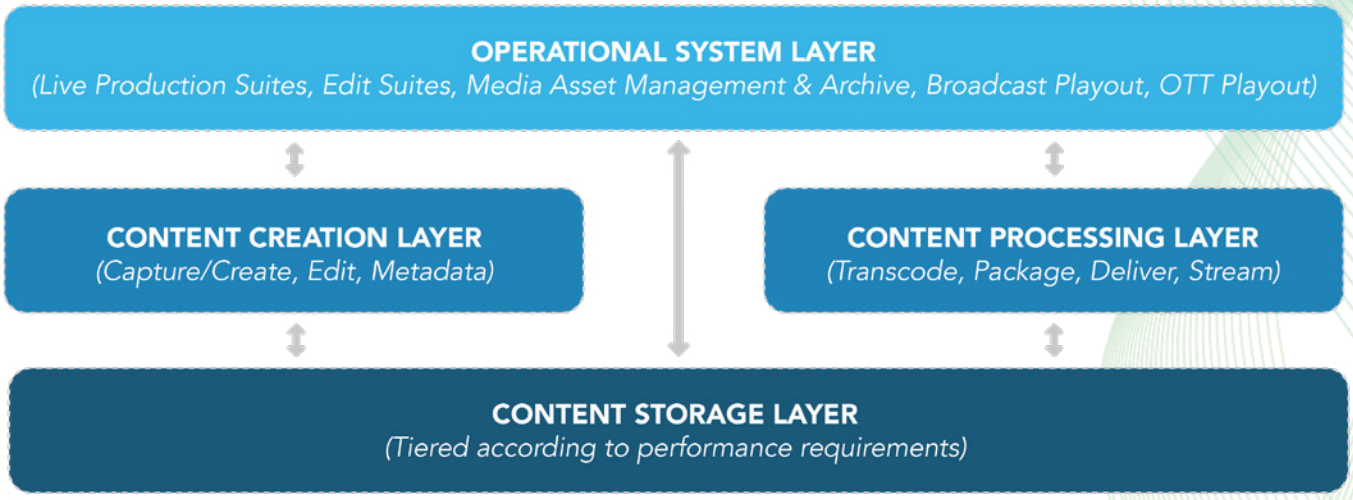
- Always accessible content for easy customer consumption and operational use
- Fast access to content for best customer experience and operational agility
- Avoidance of content migration that leads to lack of access and operational risk

FOUNDATIONS OF YOUR TECHNOLOGY

The base technology platform that can achieve these outcomes is the critical business enabler. There are various elements within the platform, including operational systems like a Media Asset Management and OTT delivery system, plus the underlying media processing applications, like transcoding or content delivery systems.

Further beneath these applications is the infrastructure that helps or hinders the efforts to achieve the 3 key outcomes. Considerations about commodity hardware, software-defined solutions, SaaS, cloud-hosted services, virtual machines, containers, bare metal deployments, and other infrastructure topics are all important. These matters are an integral part of each layer in your technology ecosystem (see below).

The underlying layer that is the foundation of the ecosystem is the storage layer. Storage interacts with each piece of the ecosystem and needs to be an always-available environment. Integration with storage, and the maintenance of that integration, is necessary for each application in each layer. As the foundation stone of a Media Operation, storage has to be the strongest part of the ecosystem. Cracks in this foundation will become evident as weaknesses in operational and financial performance.



AVOID BEING THE WEAKEST LINK

The Media industry has been gradually dealing with the shift to this future world of always-available content for operations and consumers. Like every major business shift caused by technological advances, such as the Telco industry's shift from phone service to broadband service, weaknesses are not problems until they are exposed by the new world order.

Today's Media Operations, dealing with the demands of both their business stakeholders and their consumers, have begun to expose weaknesses in their legacy storage infrastructure that until now have been largely hidden. On the other hand, the newest entrants to the Media & Entertainment industry do not have the same problems to face, as their businesses are being built on the latest technologies. Four weaknesses affecting legacy infrastructure stand out:

1. **Limited accessibility of tape-based storage.** In many ways, tape storage for a modern media business has become an operational design flaw more than a crack in the foundation. The design flaw of tape is that it is not designed for an always-accessible world. The "new-builds" of the Media industry, like Netflix and YouTube, that have been built to be "always accessible" don't have this problem. They are both built on foundations of disk-based storage. YouTube use tape but only as a 3rd-level back-up platform. Both Netflix and YouTube have had the benefit of being created in a time when tape was not the latest platform for content storage. But most Media organizations were created even before digital tape became the primary storage platform, and have ecosystems built around these tape systems.
2. **Performance limitations.** Scalability in this context refers to "storage system scalability", as opposed to the ability to simply add disks or add larger disks. Storage system scalability is an architectural and technical issue and reflects the underlying technical capabilities of a storage technology. Legacy storage systems have certain constraints that are part of their original designs. Tape is a physical storage technology with devices that read and write content, and tapes are idle until required. But tape robots and tape-drives must be used to access a tape and read/write. File-based storage technologies rely on "agents" to control file inputs and outputs. Agents are the gatekeeper to the storage, and pick and place files (much like tape systems) in specific folder locations. These system-specific access controllers become performance bottlenecks as systems grow, resulting in content taking longer to store and to retrieve.

3. **Lack of scalability means migration.**
   In file-based storage systems, the access controllers generally pre-define and then limit the total capacity of the system. When the limit is reached, the system cannot grow any further. At this point, a new and separate system has to be added to increase capacity. The normal approach is to build a new larger system alongside the original system, then transfer all content into the new system and remove the original. This is an expensive and time-consuming approach, especially as libraries reach hundreds of terabytes and more. Migrating content like this is an unwelcome element of a modern media operation.
4. **Limits to Data Protection.**
   In file-based systems, such as a NAS system, data is stored in two separate areas - the storage system itself for the content data, and the relational database for the metadata - one without the other is useless. The fact that content and the metadata that describe it are held separately creates the opportunity for disassociation, i.e. "orphaned files" that must be reconnected. Second, most data protection schemes in disk-based environments are based on RAID (Redundant Array of Independent Disks). When RAID-based systems were introduced in the late 1980s they were a significant improvement on then-current technology. Today, with larger storage arrays and disks, RAID has become a liability and a new storage method has emerged.

WHY OBJECT STORAGE WILL REPLACE FILE-BASED STORAGE IN MEDIA OPERATIONS

The latest storage technology is Object storage, a technology step-change that the new entrants to the Media industry have adopted to address the combined requirements of scalability and accessibility, as well as monetization and protection. Object storage is designed to be inherently capable of addressing these four requirements more effectively and efficiently than legacy file-based storage systems.

SCALABILITY

Object storage has been designed for the information age. Object storage design principles are focused on how to store very large quantities of data and support high growth rates, while maintaining performance of the system and avoiding the need to migrate to a new architecture. Therefore, most object storage solutions are designed to be "scale-out", not "scale-up". This means that as capacity is added to a system, extra performance is added as well. In short, the storage capacity can become as big as it can physically be, and not slow down.

As an example, object storage solutions can store billions of files. File-based storage systems generally max out at hundreds of thousands of files, and are limited by a directory structure that limits the number of files per directory, mainly because of the limits on controller architecture described earlier. If you want to store an ever-growing video library with all its metadata, supporting documents, images, audio tracks, and more, then object storage can handle it. Capacity limits for object storage systems are estimated to be in the multi-Exabyte range (i.e. thousands of Petabytes). If you are worried about the thought of migrating your current content store of tens of petabytes from one scale-up system to another, then moving to object storage could remove that concern for a very long time to come.

In terms of performance, object storage is designed to support the addition of network capacity and storage capacity, concurrently. This critical performance feature is achieved by a fundamental difference between object and file-based storage technology: object systems store files with unique identifiers, whereas file-based systems store files with file names. The end result is evident: Object systems can store and retrieve billions of unique identifiers, whereas billions of file names, simply based on file naming and storage conventions, cannot practically be stored in one system.

The unique identifiers can also be retrieved much quicker than a file-name search, partly because of the ability to locate the file without a folder trawl, and partly because there doesn't need to be an Agent to manage the process of writing and reading files to and from a specified destination point, i.e. a folder. In object storage, there is no "specified destination point" - the storage point is "in the storage" instead of "in folder 1, then within folder 10, then within folder 15, etc...". This design principle really matters to operational performance, especially when content libraries are growing and accessibility drives revenues.

In addition to be being a scale-out solution, object storage can utilize different performance levels in the commodity hardware it is using. This means that a high-performance application, such as an OTT Origin Server, and a lower-performance application, such as a media Archive, can be supported from a single storage environment. This presents a new opportunity for economies of scale and cost-efficiency as multiple applications can run in a single, unified environment. Maintenance, expansion and administration of the single system are simpler and quicker processes compared to managing a set of independent storage silos.

ACCESSIBILITY AND MONETIZATION

These two requirements generally go together. In other industries that provide products and services to the general public (like the Media industry does) the more accessible your content the more revenue it can generate. Netflix and YouTube have shown how they can attract an audience with easily accessible content, and Apple iTunes revenues continue to grow. Making content available for easy consumption requires that it resides on electronic hardware, enabling immediate access. Any disk-based storage system, including object-based and file-based, support this.

However, object and file-based storage differ in search speed. At its core, because of object's unique identifier feature, object can search and retrieve data faster, particularly in large storage systems. It is true that many Media Operations transactions will search through a database of some sort, that is designed to return a very quick response. The benefit of object storage in this case is that metadata stored in the database, used for searches, is also stored inside the object storage system, as an integral part of the object. An "object" can be thought of as the file plus the metadata. So, while search-ability may be best served by a separate high-performance database, retrieve-ability is best supported by an object storage unique identifier that includes the same metadata as the database. The unique identifier and metadata-association give object storage a performance edge over its file-based peers.

Good accessibility to content in a disk-based storage system is a benefit of both object-based and file-based solutions. Where the two technologies generally differ is in the burden of content migration. Then the "scale-up" file-based system presents a significant downside to the Media Operation. To migrate content, the Media Operation has to implement careful controls to assure content is available for live operations while also being carefully moved from library to library. This is a high-risk process. A scale-out storage solution significantly reduces the burden of migration because of its underlying design. With the natural evolution of hardware to higher performance and higher capacity, and growing library sizes, it will be inevitable to migrate regularly to newer forms of storage hardware.

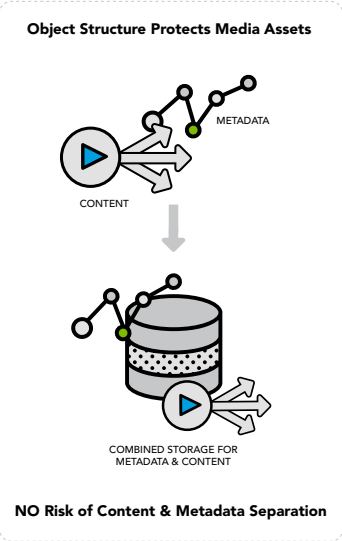
Object storage is designed to look at its hardware as a "pool" of capacity. The whole storage "cluster" is available for use by the storage system, however large it is. When hardware is changed or added, the cluster sees the new capacity available to it and distributes content accordingly per rules set by the system administrator. These specific design features of object storage mean that hardware can be changed while the storage system is in-service reading and writing data. Accessibility, retrieve-ability and performance are not affected at all. Some people refer to this as an "evergreen" solution. Object storage technology plus ever-improving commodity hardware is evergreen.

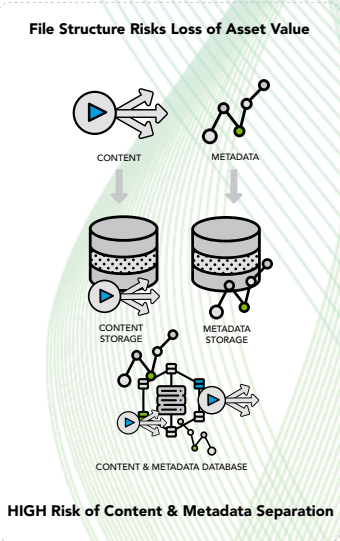
PROTECTION

The key points of scalability, accessibility and monetization point towards a critical need for business continuity measures. What good is the large and flexible storage system that is always accessible and contributing to my revenue, if data and metadata are at risk of being lost, stolen or corrupted? Media assets could literally lose value instantly if not protected.

Data protection is a large topic in its own right, and a separate paper will soon be available on this subject. There are 5 key points about protecting your business assets when considering object storage vs. file-based storage.

1. **How object storage systems store the data**
   Storage experts know all about RAID (a "Redundant Array of Independent Disks"). RAID operates at the server level, protecting the customer from data loss when hard-disks fail. Erasure-coding is effectively RAID at the cluster level - or "super-charged RAID". A cluster has multiple servers, and erasure-coding spreads data across all available servers in the cluster. From a data protection perspective this provides much greater resilience to protect from data loss in very large content libraries. RAID systems are known for overloading themselves when a disk fails and data has to be rebuilt on a new disk - this is a scary moment for storage administrators. The overload is caused by the content rebuild process consuming extra server resources while data is being written and read from the remaining disks. With erasure coding, the entire cluster absorbs the full load of a rebalance, helping reduce the stress on any one set of disks in the cluster. For large libraries running on disk-based storage, erasure coding is a best practice for protecting always-accessible data.
2. **How metadata is stored**
   Metadata adds value to stored content. Without metadata, a piece of content in a library loses its value, and recreating the metadata is an onerous task. So metadata must be closely protected. Object storage offers a key advantage when compared to file-based storage: the ability to store user defined metadata, including but not limited to video metadata, within the object itself. File-based storage systems generally store metadata separately and associate it with the content. The risk this presents is that the content and the metadata may become disassociated during content moves or system maintenance,





and "orphaned" files with unknown content can appear in the storage, creating content integrity issues. Object storage systems avoid this problem by making the object a combination of content and metadata, thereby protecting the value of the asset, and avoiding wasted time managing data in the Media Operation.

3. **How you back up metadata**
   As noted before, databases can be critical components for fast content search. But because databases can be corrupted, and sometimes lost completely, there needs to be a back-up. Of course, the back-up of the database will normally be a back-up database. But object storage offers the helpful value-add of being your database back-up because it holds the metadata along with the content, making it possible to restore both content and metadata if there is a database failure. This valuable feature is one of the protection mechanisms built in to object storage systems.
4. **How data scrubbing prevents loss**
   In tape storage systems, it is well known that tape degrades over time. So data has to be migrated from older to newer tapes to protect data. In disk-based storage, the primary issue is bit rot – the degradation over time of the electric charge of individual bits, which reduces system performance and causes files to become unreadable or lost. Some modern object storage systems utilise a proactive, automated approach to repairing bit rot known as "scrubbing". Scrubbing sweeps the system, identifying damaged or lost bits, and rebuilds primary data from supplementary data created during the erasure coding process. In the worst cases where bit rot is significant, which generally occurs in rarely accessed disks, scrubbing can trigger an entire disk to be proactively rebuilt. Scrubbing makes data protection more proactive and more efficient than before.
5. **How you assure the integrity of your data**
   Given the risks of losing content and metadata, Media Operations should continuously know their content is intact. This "inventory confirmation" task is non-trivial. Counting digital media assets at the most basic level is done by checking the database. But knowing exactly what the content is requires watching and listening to the video, a task that most content owners cannot afford. This presents a conundrum as content libraries grow. In essence, "counting" is a non-value add task. It takes time and doesn't add customer value to your product or service and is done only so that your operation can meet performance requirements with known available inventory.
   To avoid non-value adding tasks (i.e. to increase efficiency) and to confirm inventory is intact requires a technological solution. Object storage provides this solution. By storing the file with the metadata, the risk of disassociation is avoided. By scrubbing data and using erasure-coding, content is continuously protected. Of course, other business continuity best practices can be deployed in parallel, like storing additional copies and multi-site clusters, and even using tape-based back-up storage. But the primary content store, on object storage, can be more resilient and more available simultaneously.

VALUE FOR MONEY - THE ACID TEST!

A new storage solution is a significant investment for a Media business. Object storage promises excellent ROI based on its inherent capabilities to scale-out more easily, store content more securely, and be more quickly and easily accessed by internal operations and consumers. However, as business executives know very well, the benefits have to be worth it. So how else does object storage offer greater value for money than the alternatives?

1. **Using commodity IT hardware provides maximum opportunity for economies of scale.** Most object storage solutions use commodity IT hardware. Customers benefit from the continual innovation and cost-reduction from the server and disk technology industry that invests tens of billions of dollars in R&D each year. The cost of the hardware components will inevitably be as low as possible with such scale. But be careful, not all object storage solutions use commodity IT hardware, so always check.
2. **A major open-source effort has emerged to deliver Object Storage, providing an attractive alternative to proprietary solutions.** Open-source solutions like Linux have enabled technology to advance at lower cost. In storage, a major open-source effort has emerged for object storage, which provides an attractive alternative to proprietary solutions. Ceph is the largest open-source storage solution. The expectation is that Ceph will help organizations scale-out their disk-based storage infrastructure into highly-resilient, evergreen, exabyte-sized clusters. Leading organizations such as Comcast, Bloomberg, Walmart and CERN are already using Ceph at scale, helping to make it mainstream, just like Linux.
3. **Combining object storage with support for legacy file protocols (e.g. NFS, SMB) creates multi-workload capabilities.** The best object storage solutions are flexible enough to simultaneously support applications that use file, object and block protocols, and offer solutions that range from high performance to low performance, all within a single system architecture. This provides the best possible economies of scale for a storage ecosystem, and the easiest transition from legacy storage to object storage.

TCO comparisons of object storage vs. file-based storage show the biggest benefits as systems scale and organizations strive to simplify storage infrastructure. Adding extra capacity to a highly expandable system, rather than installing a completely new system in order to have extra capacity, is a significant difference in cost. Simplifying infrastructure with a unified platform, with a simple disk-based architecture under a single system management interface, removes the islands of storage that typically complicate system maintenance and operational performance. Media businesses, with the large quantities of content and the frequency of content reads and writes, stand to benefit the most from the capabilities of object storage.

THE CONCURRENT SOLUTION

Concurrent has two decades of media storage experience and now leads the industry in responding to the trends in media content growth. Concurrent's Aquari Storage is a best-in-class, modern, object storage solution, optimized for video applications and designed to support the media industry transition to petabyte-level, multi-workload, mission-critical video applications.

CONCURRENT STORAGE SOLUTION BRIEFS

- Broadcast Media Operations Storage Platform
- Archive Storage
- MAM Storage & Transcode
- Service Provider Media Storage Platform

CONCURRENT SOLUTIONS FOR MEDIA BUSINESSES

- For Content Creators and Broadcasters
- For Cable Operators and MVPDs

TO LEARN MORE ABOUT HOW CONCURRENT CAN HELP YOU TRANSITION TO OBJECT STORAGE, PLEASE VISIT WWW.CONCURRENT.COM OR CALL +1.877.978.7363