M+E JOURNAL

It's Showtime!

Innovation explodes across every workflow as technology emerges from the pandemic.

Where are you in this accelerated evolution?

DIVERSITY & INCLUSION
In the office, behind the camera, and on the screen, diversity is crucial

SECURITY
Remote productions create new security concerns, with assets under siege

SMART CONTENT
Artificial intelligence and machine learning are being applied in new, exciting ways

NEW WORKFLOWS
The cloud is delivering on its promise, powering the future of productions

21.01

SECURELY PRESERVING DIGITAL CONTENT IN A HYBRID STORAGE ARCHITECTURE

ABSTRACT: Media workflows have gone through tremendous transformation with traditional infrastructure and processes giving way to virtualized operations using cloud compute and storage. Valuable media assets, however, still need to be secured for long-term preservation. With thoughtful planning, hybrid storage architectures can be deployed so both active access and managed preservation can be achieved.

By Linda Tadic, CEO, Digital Bedrock, and Whit Jackson, VP, Technology Alliances, Media, Entertainment, Wasabi Technologies

Media workflows have gone through tremendous transformation with traditional infrastructure and processes giving way to hybrid operations blending on-premise compute and storage hardware with cloud services. Cloud storage, in particular, has assumed a far greater role in professional media operations where the data sets generated by high resolution cameras and VFX work have grown exponentially in recent years requiring additional storage capacity. But hybrid storage architectures are not just limited to the front-end production stages of media, they also play an important role on the back end where long term preservation is the goal.

MANAGED DIGITAL PRESERVATION

In the analog film and videotape-based era, media companies knew how to manage their assets. Physical film and tape masters were locked in vaults, and industry guidelines outlined required actions to ensure the safety of the company's intellectual property. Studios retained experienced archives staff to keep their valued assets alive through film restorations, and reformatting videotapes from obsolete video formats to "new" video formats — that then became obsolete.

Digital assets still need to be managed, restored and reformatted as it was before with

AS WITH ANYTHING in the media world, change and technological evolution is an ongoing reality. Yet the need to preserve valued digital content, in any medium, is of critical importance for perspective and appreciation. As the archival adage says, 'Preservation without access is pointless.'

film and tape. Nothing has changed as far as the core concepts in preservation work. However, the technologies to manage and perform preservation actions on valued content have evolved, requiring a different kind of diligence. Digital preservation requires monitoring and mitigating the digital object obsolescence itself, the health of the data, and storage media migrations. Digital preservation must be managed — storage is just one part of an overall digital preservation strategy.

Digital object obsolescence. Digital formats can become just as obsolete as the familiar videotape formats from the past. Revising and updating computing and storage devices is a necessity as technologies improve, but legacy content manifested as digital objects become endangered when the software, operating systems, and hardware required to render a file become unsupported as new technologies are developed. In order to monitor obsolescence vulnerabilities, content owners must know the technical characteristics of the digital objects and files in their collections, and what software, operating systems, and hardware are required to keep that content usable. Digital object obsolescence is separate from storage media obsolescence (for example, data tape, hard drives, and file system obsolescence).

Bit health and fixity checks. Digital objects are bitstreams contained in formats. A central principle in digital preservation is to verify that the bits are still "fixed" and unchanged over time. Readers might be familiar with the term "checksum," and if so, most familiar with the md5 checksum. Broadly speaking, a "checksum" or "hash" is an alphanumeric value generated by a mathematical algorithm. The output, or hash value, should be unique to a particular file. There are several types of hash algorithms, with SHA-512 considered the most secure in common use (e.g., its complexity makes it more difficult to hack or spoof). Regardless of the algorithm used, digital content should be checked according to a schedule over time to verify that the file has not changed. In digital preservation language, this process is called a "fixity check." If a file is retrieved in five years and a "fixity check" run against its original checksum or hash, it should be the same. If it is not, then something has changed in the file.

How often fixity checks are run depends on how the files are stored. This article does not have space to delve into the different methodologies of fixity checks in object storage in the cloud, versus checks on standard file-based server storage, or checking files on data tape. The main point to stress is that checksums/hashes — the unique "fingerprint" — should be generated the moment when a file/object is finalized, and then verified according to a schedule over time.

Media migration. Data is stored on something, whether spinning disks, solid state drives, or data tape such as LTO. Locking storage media in vaults behind a series of physical barriers and several layers of authentication is a strong method to protect digital content from unauthorized use. Of course, the offline storage media itself must be migrated over time as storage media ages and becomes obsolete, with fixity checks

---

***Whit Jackson** is VP of technology alliances for media and entertainment at Wasabi Technologies. He is driving the adoption of Wasabi hot cloud storage in media and entertainment creation and distribution, and throughout his career has worked at the forefront of innovation in TV and film production and distribution. wjackson@wasabi.com @wasabi_cloud.*

***Linda Tadic**, CEO of Digital Bedrock, is a thought leader in media and digital preservation and metadata. She has more than 30 years' experience in leading preservation, metadata and digital production operations at studios, archives and cultural heritage organizations. ltadic@digitalbedrock.com @LindaTadic*

performed over time and when the migration occurs. Migration — of both storage media and file formats — is a fact of digital life and planning for it should be part of a digital preservation strategy. Successful storage media migrations will be helped if the files' checksums were verified after writing to offline media prior to storage, and open file systems are used in writing the data to storage so the files are not locked in any proprietary software.

LEVERAGING CLOUD STORAGE FOR AGGREGATION, REDUNDANCY AND REMOTE ACCESS

In general, cloud storage offers a number of benefits for media organizations. First, the cloud provides on-demand scalability. When a new production project kicks off or a new media library is acquired that stresses the capacity of existing storage infrastructure, the cloud provides immediately available and limitless capacity. Second, cloud storage also provides financial flexibility and reduces capital expenditure outlays. You pay for what you use on an operating expense basis, which is very helpful when projects ebb and flow and workloads are constantly changing. There's no hardware to procure, install, power and maintain. Lastly, cloud storage provides remote accessibility to content from virtually any location with a decent network connection. And if using a hot cloud storage service, all content is immediately retrievable for review and repurposing.

In digital preservation, cloud storage also serves as an effective complement to the storage of offline media such as LTO tapes in highly secure vaults. Unlike the days of collecting up film reels and video tape cartridges for review and restoration before archiving, digital archiving involves the collection of data files. At this stage, cloud storage becomes a very convenient location to aggregate all the material considered for long term preservation. Digital files can be sent securely from various locations over the Internet or private network connections to a common cloud account and then be reviewed and organized remotely before committing assets to preservation processing and offline storage.

The cloud account can also serve as a secondary copy of the offline archive that can be readily accessed for review, clipping or re-release. Depending on the circumstance, a new project could be initiated using the cloud copy or by restoring the "truth" copy from preservation.

CONCLUSION

As with anything in the media world, change and technological evolution is an ongoing reality. Yet the need to preserve valued digital content, in any medium, is of critical importance for perspective and appreciation. As the archival adage says, "Preservation without access is pointless." With thoughtful planning, hybrid storage architectures can be deployed so both active access and long-term managed preservation can be achieved to protect fragile digital content for future use. ⊞

DIGITAL BEDROCK

PROTECT YOUR DIGITAL CONTENT WITH OUR SECURE, MANAGED DIGITAL PRESERVATION SERVICES

We are your staff and infrastructure:

NO software licenses

NO annual subscription fees

NO staff training

NO large-scale hardware or infrastructure investments

NO minimum or maximum storage requirements

Easy exit

Digital Bedrock preserves content through:

Detailed file analysis to catch hidden errors before long-term preservation

Extensive metadata extraction and indexing

Active file format obsolescence monitoring

Scheduled SHA-512 fixity checks to verify bit health

Three copies geographically dispersed in secure, offline storage

Fast direct connection to major cloud storage providers

Combining intelligent object storage search with offline security

Digital Bedrock supports all complex and proprietary formats.

INFO@DIGITALBEDROCK.COM | 888.938.7386 | WWW.DIGITALBEDROCK.COM