Storage for Next Generation Workflows

Overcoming latency issues with memory first storage.

Introduction

Digital media projects are increasing in size due to higher resolutions and frame rates, higher dynamic range (HDR) content and an increase in the number of cameras used for each project. Video professionals also tend to capture more content with today's digital recording than they ever did with film and older recording technologies. The number of professional video projects has increased to feed content to more distribution channels and the timelines for completing projects are more and more constrained.

Single video projects can be several petabytes (PB) in size today. This storage requirement may swell to 100's of PB of data for a single video project within the next 10 years. Exabyte workflows are in the near future for some video projects and potentially for production houses that manage multiple large video projects. Media integrity and effective replication are important to retain this expensive captured content

Multiple high resolution, high frame rate, high dynamic range and multi-camera projects generates a large number of media files.   In addition, the sizes of these individual files are large and growing. Higher performance, (both higher data rate and lower latency) is needed from the storage infrastructure that supports these workflows to enable digital media professionals to do their jobs.

Hardware can solve some of the issues of working with large media, but software will play an increasing role. New software tools, increasingly using various AI techniques, are helping media professionals keep track of and more effectively use their stored content.

At the same time, the demand for more storage must be matched against storage budgets that are flat or only increasing by a small amount annually. The industry needs advanced tools to meet these technical requirements while meeting industry spending constraints.

Modern Post Production

Non-linear editing (NLE) is generally done with un-compressed or at most slightly compressed source content since heavy compression can lead to loss of resolution and can cause timing problems.

Figure 1 is a schematic of a non-linear editing station showing connection to shared on-line (or real time) storage via a SAN or NAS (or a cloud-based object storage system) via a host bus adapter (HBA or xGbE card).   Modern editing stations may use a combination of local (direct attached), networked and remote storage.
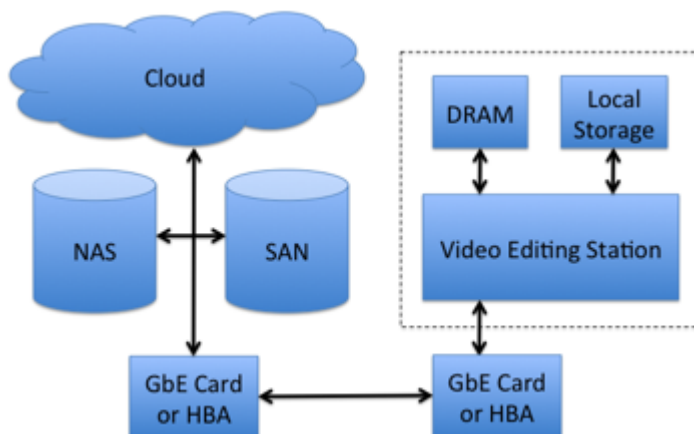
Figure 1. Professional Non-Linear Editing Model System[1]

Latency and Its Effects

Latency is the time it takes from when a command is issued, such as to read a video file, and when that content is received at a workstation and ready for professional processing. Latency is affected by other performance factors such as bandwidth and throughput, but certain aspects of latency can be controlled within the storage solution itself. Latency is actually a fairly complex notion, encompassing many factors, including data transport distance (time to transmit data across a transport medium) along with any electrical and/or mechanical processes that the storage solution must carry out in order to fulfill the overall command. If the latency of the storage solution is too high it will impede the work of the video professional and impact productivity.

For the media and entertainment industry, high latency equates to an experiential deficiency. Post-production shops have many concurrent users all accessing and working with project data collaboratively, and file access times can be impacted not only by the number of concurrent users but also by the type of storage architecture (NAS or SAN). High latency complicates the loading of project files, scrubbing back and forth through the project, and overall playback quality.   For post-production workflows, this can be a real problem.

For instance, for data located in a remote data center (e.g., in the cloud), the transport time can be a significant part of the overall system latency. This is why most use of media content in the cloud involves downloading or uploading content to be processed locally, rather than real time processing in the cloud (although streaming of content is the cloud and some lighter editing using lower resolution proxies processed locally, can be done successfully using cloud-based content). The latency inherent in direct attached or local network storage is the bigger issue for most post production facilities. Traditional spinning hard drives, for example, take time to spin platters and access data from or write data to the physical storage medium. Solid-state drives, while having higher performance than hard disk drives, take longer to write data than to read it.

Logical components such as software interfaces and operating systems also add to overall latency. And unlike transport variabilities, these latency-inducing factors are entirely under the control of storage solutions vendors. The right technology will reduce latency at the component level and have a noticeable effect on the end user experience.

OpenDrives Products

OpenDrives works closely with its customers in the media & entertainment, gaming and healthcare industries to provide storage solutions that greatly reduce latency and improve response times and the overall user experience. As mentioned earlier, this is a necessity in post-production workflows, where many users are accessing project files simultaneously, and smooth high-quality playback is key to the overall workflow. Because of its deep institutional knowledge of the industry, Open-Drives has implemented architectural features to overcome the latency factor. OpenDrives' storage solutions reflect a "memory first" approach. What this means is, implementing control logic to ensure the use of high-speed memory over slower read/write operations performed by mechanical drives. OpenDrives' privileges memory caching and high-speed secondary caching to avoid latency-inducing mechanical operations.

For example, when data is about to be written to storage, smart logic with the OpenDrives' operating system ensures that the data lands first in memory prior to being committed to disk. To add to that, data remains in memory for a certain period of time in case the data is frequently accessed. In that case, it is read from memory, not disk. This vastly reduces latency within the overall transaction.

This same approach is used for read operations. When data is read from either an HDD or SSD disk within an OpenDrives' solution, our software puts that data into memory at the moment it's read, and it stays there. That way, if the application requests that data again, it's read directly from the much faster memory cache. To reduce latency even more, OpenDrives' uses predictive algorithms to "pre-fetch" data. When a file sequence is being read, the software preloads the next file into the faster access tier so it's ready even before the

The operating system is constantly evaluating the files in memory, and as the files become less active over time, they eventually are retired from cache to make room for new pre-fetched data. All of this activity is embedded at the operating system level so it is completely transparent to the user, and does not depend on a specific type of drive (flash or HDD).

OpenDrives leverages the low latency capabilities of their storage systems with software that involves more than just basic machine learning or MAM software integrations that other storage providers are calling AI.

OpenDrives scalable storage solutions for high-end media workflows are their Apex, Summit, Ridgeview and Pinnacle systems. These storage systems are all flash, hybrid flash and HDD, and all HDD storage systems. The company manages these systems with their ATLAS OnmiManagement Software. All OpenDrives systems include:

- Industry standard networking protocols
- No per-seat software licensing fees
- Flexible Ethernet architecture that requires no special connection software or hardware
- An easy-to-use management interface (ATLAS)
- Exclusive bandwidth monitoring and throttling engine
- Fault tolerant data pools
- Innovate IOPS balancing engine so one client operation doesn't overwhelm all others
- Advanced metadata management
- Inline lossless compression and de-duplication
- Fully check-summed file system providing unmatched data integrity and protection
- Reduced energy costs, physical footprint, heat production and support costs
- Greater data portability
- Best-of-breed hardware engineering for long-term resilience and reliability

Making Gone Girl

Shooting Gone Girl in 6K meant Fox and David Fincher needed a storage system that could meet demanding requirements. Playback of a single stream of 6K DPX files requires 1.6 gigabytes/second, performance no traditional centralized storage system could meet. Additionally, they needed to be able to apply real-time transformations on reframes, as cropping and reimporting would slow down their aggressive workflow.

As a further guarantee of data integrity, OpenDrives stores data redundantly, both physically and logically. If the checksum for a given block does not match, the parity data is used to repair the bad block of data. What this means is that our solutions rewrite the incorrect (or corrupt) block and checksum with correct data and values. This means that Open-Drives storage solutions are self-healing due to this proprietary logic within our operating system.

Fox required concurrent editing, VFX, and compositing of different shots without transferring back and forth between local workstation storage and the networked share. This meant the entire film had to be kept online and centralized throughout the creative process. To accommodate various disparate workflows and the comfort level of the creatives, the storage system had to work seamlessly with Mac and Windows devices.

OpenDrives' Apex delivered the perfect solution. The only system that could deliver the playback requirements, Apex's all-flash system allowed editors and visual effects artists to work on multiple projects on a centralized shared infrastructure.

Delivering coherent permissions and accessibility from both OSX and Windows workstations, the OpenDrives Avalanche NAS was self-contained in a single 4 Rack Unit controller that simplified deployment, maintenance and footprint.

VFX shots and composites were completed in-house and parallel to editing. OpenDrives' creative fluidity allowed project files to be opened, recognized and adjusted on the fly. Apex ensured smooth scrubbing and playback, faster exporting and transcoding, and reincorporated completed composites 200% faster than traditional render and replace workflows.

In making Gone Girl, David Fincher and Fox pushed technologies and workflows to their limits. The OpenDrives Avalanche was the only storage solution that could deliver both the performance demands and ensure the creatives were provided the toolset they needed focus on their process.

Corgan MediaLab

Corgan MediaLab is the media wing of the Corgan architectural and design firm. The MediaLab engages in many different forms of video production and post-production, from animation and motion graphics to VFX and 3D architectural renderings.

Corgan MediaLab had experienced difficulties with their creative workflow and pinpointed the storage infrastructure as the bottleneck. In addition to a non-linear workflow, versioning was non-existent at the time. Load time for small videos and playback would lock up the network.

OpenDrives took a consultative approach with the professionals at Corgan MediaLab to support their varied production workflows, comprising 3D, VFX, film services and a host of other high-intensity workflows. The OpenDrives production solution (Summit) was augmented with a backup and archiving solution (Ridgeview) to replace a tape-based workflow and help facilitate reuse of footage.

By selecting OpenDrives, Corgan MediaLab was able to build a real-time workflow for their mixed media environment providing the team members with a shared environment for collaboration across multiple applications and workstations. OpenDrives delivered a solid technology solution and support, resulting in the highest level of customer satisfaction.

Conclusions

Latency is one of the most critical problems for post-production on high resolution, high frame rate, high dynamic range video workflows. Latency can be kept as low as possible using local network storage with intelligent movement of data from storage devices to and from high-speed memory. Intelligent movement of data and overall data management allows the most cost-effective digital storage solutions to serve the media and entertainment industry. OpenDrives provides storage solutions focused on minimizing video latency during post-production.

As the size of video projects increases, so does the number and size of the video files. Intelligent management of these assets is a key element of cost effective workflows. Today, products such as OpenDrives' OmniManagement software helps media professionals maintain and find their data, but in the near future, AI technology, intelligently focused on the needs of media professionals, will make software management and organization of data even more powerful.

About the author

Tom Coughlin, President, Coughlin Associates is a widely respected digital storage analyst as well as business and technology consultant. He has over 37 years in the data storage industry with multiple engineering and management positions at high profile companies.

Dr. Coughlin has many publications and six patents to his credit. Tom is also the author of Digital Storage in Consumer Electronics: The Essential Guide, which is in its second edition with Springer. Coughlin Associates provides market and technology analysis as well as Data Storage Technical and Business Consulting services. Tom publishes the Digital Storage Technology Newsletter, the Media and Entertainment Storage Report, the Emerging Non-Volatile Memory Report and other industry reports. Tom is also a regular contributor on digital storage for Forbes.com and other blogs.

Tom is active with SMPTE, SNIA, the IEEE (he is President-elect of IEEE_USA and active in the Consumer Electronics Society where he is chairman of the Future Directions Committee) and other professional organizations. Tom is the founder and organizer of the *Annual Storage Visions Conference*, a partner to the International Consumer Electronics Show, as well as the *Creative Storage Conference*. He was the general chairman of the annual Flash Memory Summit, the world's largest independent storage event for 10 years. He is a Fellow of the IEEE and a member of the Consultants Network of Silicon Valley (CNSV). For more information on Tom Coughlin and his publications go to *www.tomcoughlin.com*.