

File Based Storage for Broadcast Workflows

Thomas M. Coughlin
Coughlin Associates

Sponsored by



Introduction

These are indeed interesting times. Concurrent with the wide adoption of file-based workflows, the resolution of video content is increasing while the number of formats and sources of content are expanding dramatically. At the same time, broadcast budget constraints are more severe than they have ever been. If you are facing paradoxical conditions like these, you need to know about the technologies available today providing the high performance file storage your creative processes need at a cost far below that required by earlier generation products.

This paper explores current pain points in the broadcast workflow and how the use of scalable, high performance Network-Attached Storage (NAS) can reduce this pain. You will find out how high performance scalable NAS with a single file system provides solutions to many problems in today's file-based broadcast workflows. You will also get some insights on where broadcast workflows may evolve to accommodate the mutual needs of higher performance, easier content management and constrained equipment and software budgets.

Pain Points in Broadcast Workflows

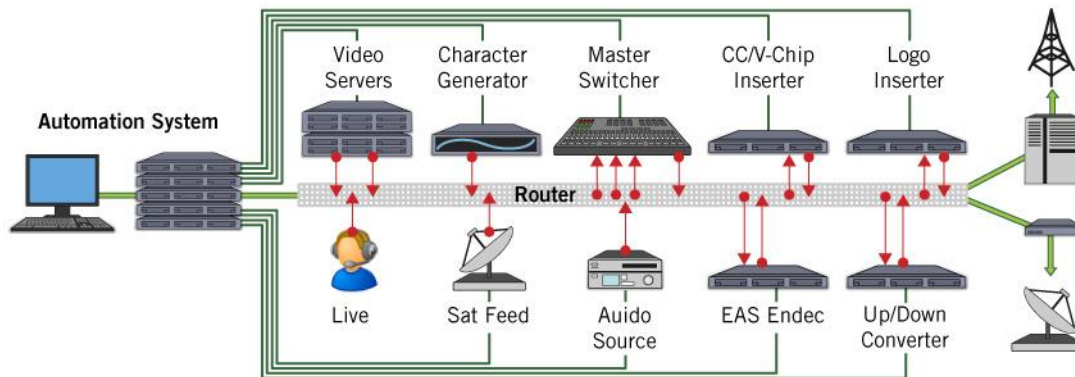
Figure 1 shows the functional complexity of a modern file-based broadcasting environment with its various operations from media management, editing, processing and routing, through release to the various distribution channels and to the actual channel qualification and use. In this paper we will focus on the parts of these processes relying on digital storage and in particular on those elements where new file-based storage solutions can help deal with the the challenges and reap the advantages offered by modern broadcast technology.

Figure 1. Technical Operations for Broadcast

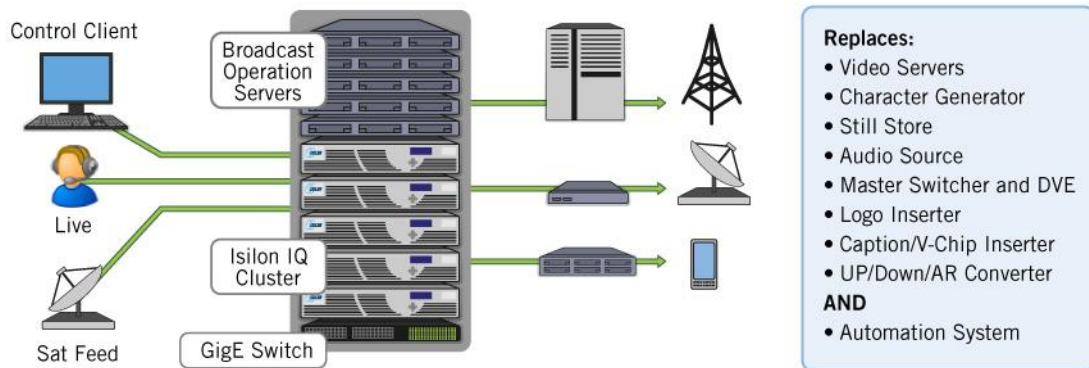
<i>Media Management</i>	<i>Networking & Editing</i>	<i>Core Processing</i>	<i>Channel Release</i>	<i>Media Transport</i>	<i>Transmission</i>
<i>Ingest</i>	Editing	Processing and Routing	Automation Servers	Test & Management	Networking Transmitters
<i>Digital asset management</i>	Graphics and Branding	Distribution	Graphics & Branding	Video Networking	Mutli-channel Headends
<i>Media movement</i>	Newsroom Integration	Compression	Master Control	A/V Networking	Mobile TV
		Network Control		IPTV	

Software based management of broadcasting operations is achieved today with greater efficiency and at lower cost than was possible with more complex proprietary hardware systems only a few years ago. Innovations in management software combined with the use of commodity computing hardware have resulted in lower up-front capital expenditures as well as lower operating costs due to increased management efficiency and the use of modular scale-out architectures. **Figures 2** and **3** are graphical representations of a video hardware vs. file-based software broadcast workflow.

**Figure 2. Hardware Based Broadcast Workflow
Multiple Proprietary Hardware Components**



**Figure 3. Software Based Broadcast Workflow
Single Rack with Commodity Hardware and Software Modules**



Major benefits of implementing modular software-based broadcast solutions include a truly concurrent non-linear approach to media access, substantially better agility to offer new formats, functions or customized workflows, significant cost reductions (by moving away from proprietary hardware) and easier management and scalability with the inherently modular architecture.

With the move to a more flexible and agile application infrastructure, the need for a more scalable storage infrastructure becomes necessary, otherwise the inefficiencies are simply moved to the storage end of the workflow with no benefits realized from the overall architecture. In order to understand the storage challenge, let's review the I/O access points between application and storage in a broadcast workflow.

Despite the many specific steps going from content capture to playout, there are three areas to address in order to alleviate key pain points. The first operation we will examine is content acquisition, or ingest, which may involve content in many different forms and formats. This content can come directly from digital cameras, downloads from a central distribution point or satellite feed or even from the digital capture of analog content. At a national or network level there may be a great many different content formats involved. At a regional and local level these formats tend to be more limited but still represent a challenge. Some content acquisition pain points are the cost of storage media on the capture devices and, most importantly, the time required to ingest content.

The second major step in content creation and distribution involves various types of content processing such as editing, graphics generation, closed captioning and appropriate metadata manipulation such as indexing to enable advertisement insertions and media management functions. Generally these operations are referred to as post production. Some pain points in the post production process are the cost of real-time (high throughput, high QoS, low latency) storage needed for creative work and the scaling of that storage to meet the needs of evolving content formats and resolutions, especially for a multi-seat facility. In particular, bandwidth and management of real-time

and near-line content in post production to support server process speeds can be a significant issue.

The third function is actual distribution of the prepared content and formats to the appropriate distribution channels. A significant amount of processing is needed to transcode potentially massive amounts of content into formats appropriate for various distribution technologies such as over the air broadcast, cable or satellite feeds, streaming or downloaded internet distribution and mobile phone distribution. Pain points in content distribution are the cost of high performance storage for centralized and edge content delivery, the bandwidth needed to handle multiple channels of simultaneous content distribution, format transcoding speeds and providing a large enough library of content to satisfy the increasing demand for “long tail” content.

While optimizing the utilization of storage within each of the distinct processes outlined above, a key challenge is reducing the time it takes to move content across these processes. This is especially critical in news operations where the time to air from the occurrence of the live event to playout can have a major impact on channel ratings.

Increasing Resolutions Drives Up Storage Capacity and Performance Requirements

In each operation in the broadcast and electronic distribution workflow there is a significant amount of digital storage capacity and performance required. Digital storage plays a pivotal role in all elements of capture, ingest, editing and distribution, and the desired characteristics of the digital storage used in these different operations can be very different.

As the broadcast and electronic distribution industry moves to higher definition content (HD today but even higher resolution and 3D content in the future) it requires ever larger content files, requiring larger storage volumes. At the same time, the access to this content must maintain its required performance and quality. It must also accommodate multiple (and new) formats when required. The requirements of higher resolution content make the pain points in file-based workflows even more acute.

As an example of the media storage demands for higher resolution HD production, the 2008 Olympics covered by NBC in China had the following statistics:

- About 1 PB of Sony XD Cam media used in 180 stand alone cameras
- About 400 TB on Omneon Media Servers in China and New York
- About 260 TB on Avid storage associated with NLE equipment and ISIS software
- About 250 TB on Isilon network attached storage servers for content delivery

DAS

Three basic types of mass storage architectures are used in media creation and distribution. The first is Direct Attached Storage (DAS) in which a storage device is directly connected to a workstation or other client, but the content in that storage device cannot be directly shared with other clients – making efficient workflows hard to implement. Direct attached storage with a fast interface can provide fast data access since the device only supports the system it is attached to. This is why DAS has remained popular in the post-production environment. That said, faster bandwidth data networks combined with the improved efficiencies and productivity offered by shared network storage are reducing the use of DAS in broadcast workflows.

NAS

Data on a direct attached device cannot be shared with others for collaborative work and it is difficult to coordinate and manage. Network Attached Storage (NAS) and Storage Area Network (SAN) storage – provided a shared SAN file system has been implemented - allows the stored content to be shared between multiple clients. With the increasing performance of networked storage and in particular, cost-effective Ethernet-based NAS, the utilization of DAS for fast data access in post-production environments is decreasing. **Figure 4** shows our expectations for networked storage to increase at the expense of DAS for post-production processes.

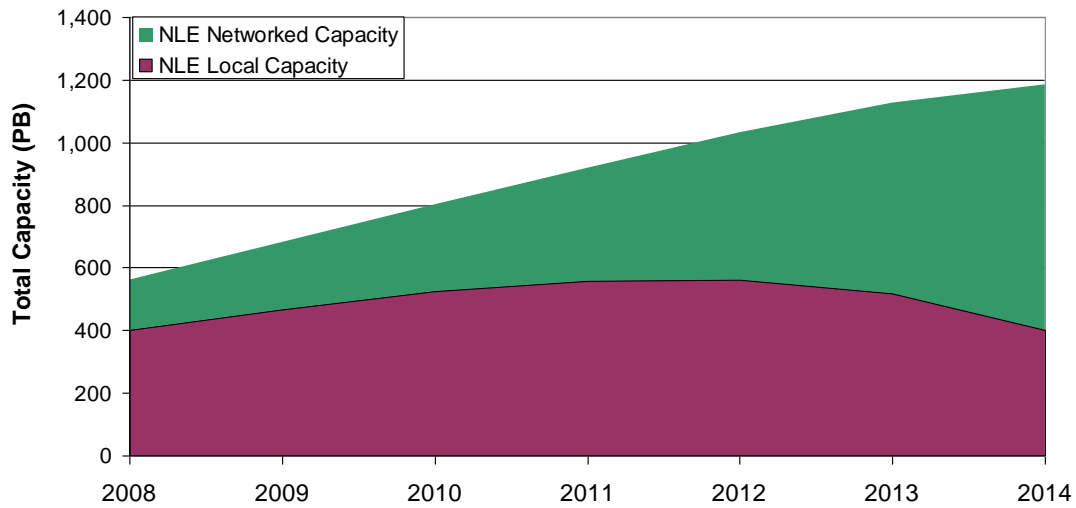
NAS storage devices provide direct file access to content. To a client, a NAS storage system looks like another hard disk drive that is connected to his or her system. Because it offers direct file access to content without an intermediary file system, NAS storage has become a very important element supporting today's file-based workflows.

SAN

In the case of a pure SAN, the storage system uses a specialized high speed networking technology (such as Fibre Channel or 1-10 Gigabit Ethernet) to provide access to data at the block level, similar to the actual storage on a hard disk drive. For a SAN, a 3rd-party file system must be located on each client system accessing the data on the storage network using specialized network adaptors.

In the past, NAS has been commonly used as near-line storage where content is kept in a secondary lower cost, slower access storage tier, while SAN was often used in expensive, very fast, real-time content systems. However, due to the increase in performance enabled by economical 1 Gb and 10 GB Ethernet networks vs. a specialized Fibre Channel network (used in traditional SANs), NAS storage is being adopted as primary storage by many content creators. A NAS system that can scale to meet performance and storage needs while providing a single file system is even more attractive to professional content users because it provides ease of use, ease of management and dramatically improved scalability. Figures 2 and 3 show the benefits of efficiency and scaling to a broadcast distribution system.

Figure 4. Growth of Networked and Direct Attached Storage for Post-Production



Broadcast and other forms of digital content distribution use compressed formats in order to make the best use of limited bandwidth resources. Even with compression, however, Video on Demand (VOD) and other distribution technologies will require larger amounts of storage as the digital content resolution increases.

Relieving the Pain in Digital Workflows

There are many ways to improve the performance of content creation and delivery systems. Following are some common approaches with comments about the strengths and weaknesses of each.

- Use of file based workflows: In modern broadcast facilities, content is a digital file from capture through editing and delivery. With digital content, file-based workflows using storage networks to share and manage content facilities see significant reductions in their costs of production as well as improvements in the quality and speed of production work. With a unified file system, all of this content can be easily accessed and centrally managed under a single namespace.
- Overprovisioning of DAS: Direct attached storage used in this industry has very fast interconnection speeds allowing fast access to stored content for a client workstation. Thus, raw performance can be improved by adding more high speed DAS storage. However, with this approach content may need to be replicated on each workstation's DAS storage and centralized management and sharing of the content for collaborative work becomes much more difficult. Unlike network storage, DAS does not facilitate collaborative work.
- Improving the speed of networks: A facility can go to 10 GB Ethernet networks or TCP/IP Offload Engines (TOEs) can be added to all network interface cards

(NICs) in each existing server. Approaches like these can solve server network I/O problems but does not solve the storage I/O problem.

- Improving the speed of storage systems used for content delivery: Highly scalable storage systems (such as those offered by Isilon) allow incremental storage and performance improvements as needed. This significantly increases access speed to content. Combined with high speed real-time storage systems using grid or blade-based storage or solid state storage for edge delivery or caching, significant improvements in content accessibility and delivery can be achieved.

Storage Scaling and Efficiency in the Electronic Distribution Market

Facing continuous pressures to deliver more with ever shrinking budgets, media and entertainment managers look closely at the initial purchase price as well as the operating budget for digital storage and other key technology investments. Solutions that can scale easily to match required storage capacities and performance and which help to reduce operating costs will be looked on most favorably.

Thus in many industries, virtualization and deduplication are being implemented as ways to improve the efficiency of digital storage systems. Using SATA drive systems for bulk content storage rather than more expensive but higher performance SAS or Fibre Channel drives are additional ways to reduce capital and operating costs. In this way, SAS or Fibre Channel drives can be used for the fastest operations while supported by a SATA array of lower cost storage for bulk content storage.

As pointed out before, in many cases media and distribution content have different requirements compared to traditional IT storage. Virtualization and deduplication in the traditional sense will not work for active media since performance requirements are, in general, not compatible with these technologies. As shown earlier, specialized storage systems that can provide good quality of service and continuous non-interrupted high resolution content delivery are needed in post-production and content delivery.

Scale-out NAS Storage Provides Solutions to a Customer's Problems

Discovery Communications, the world's leading broadcaster of non-fiction media, has implemented advanced network storage technology using Isilon Systems' scale-out NAS storage for domestic and some international media production. Note that like many modern content companies, much of the network content is produced by contractors. Creating a centralized process for working with content in multiple formats and from multiple sources can be challenging.

In Discovery's domestic post-production facility they have 60 concurrent AVID users across 3 shifts (180 users daily) for video editing on AVID, some editing for DVD and VOD and transcoding of content into other media formats to support all the different distribution channels. The company digitizes 64 hours a day of 145 Mbps AVID codec content (4 TB/day) and maintains a high resolution content repository (an active archive) for all US domestic broadcast content as well as some international content on Isilon storage systems. In addition to high resolution master copies, the Isilon storage systems provide storage for multiple content formats. DVD authoring content is also stored on these systems using SONIC DVD authoring tools.

The flexibility of Isilon Systems scalable single file system NAS storage allowed easy scaling as required by the growth of the business and thus a most cost effective capital deployment. Discovery started with 2 small Isilon clusters and as they expanded their needs they consolidated both of these clusters into larger 51-node cluster adding later generation X-Series clusters with total of 17 12000x series nodes and several accelerator nodes to add 10GigE connectivity. The result of this storage expansion was a total of 71 Isilon IQ scale-out NAS nodes managing 250 TB of storage capacity consolidated to 3 storage clusters.

These storage clusters allow a more consolidated environment where a single namespace is used across the different elements of the company's workflow. Instead of copying data across local storage devices, data can be accessed more efficiently within the Isilon clusters. Having a centralized storage pool in a scale-out NAS storage system is also much easier to manage than a distributed pool of local storage devices. This is important as content resolution and capacity increases. Average file sizes are 20GB per file for SD and up to 70GB for an hour of HD media. 3-D and higher resolution formats will swell these requirements even more in years to come.

In addition to the easy deployment of new storage as needed, Discovery uses the storage system quota management application to manage space consumption. 10TB quotas on all the storage shares are used to thinly-provision the storage presented to the end users and trigger alarms when a workflow starts to approach 80% of their provisioned storage. When system administrators receive such alerts, they either reallocate quotas based on actual utilization, or add physical storage on the fly during normal business hours in anticipation of increasing storage

"More often than not, we are presented with a new workflow problem the same month or often the same week as the live date.

We rely heavily on the ability to independently scale virtualized storage and bandwidth in a 24/7 broadcast environment without the need for downtime or weekend work."

Robert McNeal
Manager - Broadcast Network
Engineering
Discovery Communications
Technical Center

requirements to meet workflow schedules. Workflow, bandwidth, and storage requirements are constantly changing as producers move from project to project.

Discovery must continuously shape its business to keep up with the changes in media content acquisition, processing, and the expanding array of content delivery platforms (iTunes, Xbox, PSP, Amazon, etc.). As an example, storage shares can fluctuate by tens of TB on a weekly basis. In addition, and as is the case for all companies managing tight IT budgets, companies such as Discovery need to control their management time and consequent expenses for their content storage network. With the Isilon scale-out NAS architecture, they have reduced weekly management time to only a few hours per week without sacrificing creativity and productivity.

Future Storage Trends for the Media and Entertainment Market

Digital storage capacity requirements are driven by the ever increasing creation of digital content as well as digitization of extensive libraries of analog content. Much of this digital media content is kept in on-line libraries rather than simply in cold archives - providing much faster access and monetization of this content.

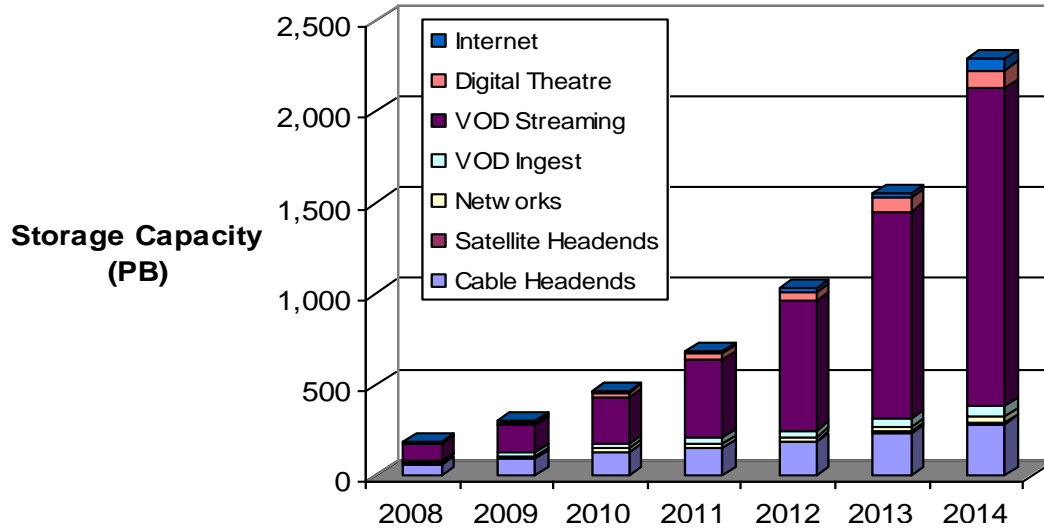
Digital content is finding more and more distribution channels. **Figure 5** shows that total storage capacity for all channels of professional media content distribution will increase over 9 times in the next 5 years. A media and entertainment professional survey conducted by Coughlin Associates showed some interesting results on content distribution trends including:

- Average hours on central content delivery system was about 200 hours (varied widely) with about 150 hours ingested monthly
-
- About 20% used flash memory on their edge servers

More content, with higher resolution and more distribution channels will depend upon creative storage approaches, such as those provided by scale-out NAS from Isilon Systems.

In the future, we might expect even tighter coupling of content and storage. More detailed and automatically generated metadata could be captured and kept with the content essence as an object recognized by the storage system itself. This could lead to increasingly efficient and effective object-based workflows evolved from today's file-based workflow.

Figure 5. Storage Capacity Growth for Digital Media Content Distribution



In Conclusion

Today's digital, file-based workflows pose increasing challenges to traditional storage. Flexible, scalable, low cost, high performance storage systems enable expanding digital storage needs for file-based workflows. The sheer size of professional content libraries requires new levels of management and control of both initial capital costs as well as on-going operational costs. New storage architectures and devices are being implemented in the professional media industry to enable the acquisition, preservation and use of this vast and expanding content library. A scalable NAS storage system offering a single namespace and scalable storage capacity and performance can increase the efficiency while decreasing the costs of high resolution workflows.

The future holds even greater promise as evolving and more intelligent commodity-based hardware with scalable content aware storage systems make handling the steadily increasing resolution, formats and quantity of content used in broadcast environments cost effective, efficient and, ultimately, profitable.

Sources

Material from these sources was used in this white paper.

- **2009 Digital Storage for Media and Entertainment Report**, Coughlin Associates
- **2009 Survey on Storage for Media and Entertainment**, Coughlin Associates
- For more information on these reports please go to www.tomcoughlin.com (tech papers page)

About the Author:



Tom Coughlin, President, Coughlin Associates has been working for over 25 years in the data storage industry at companies such as Ampex, Polaroid, Seagate, Maxtor, Micropolis, Syquest, 3M and other companies. He has over 60 publications and 6 patents to his credit. Tom is active with IDEMA, the IEEE Magnetics Society, IEEE Consumer Electronics Society, SMPTE and other professional organizations. He is the founder and organizer of the Annual Storage Visions Conference (before the International CES) as well as the Creative Storage Conference (before the annual NAB show).

Coughlin Associates provides market and technology analysis as well as Data Storage Technical Consulting services. Dr. Coughlin is the author of Digital Storage in Consumer Electronics: The Essential Guide, published by a division of Elsevier. Coughlin Associates publishes reports on digital storage in professional media and entertainment as well as consumer electronics. For more information go to www.tomcoughlin.com.