



The Essentials of Networked Media

Al Kovalick, CTO
Broadcast and Professional Division
Pinnacle Systems

Abstract

Networked Media is revolutionizing the way audio/video production and distribution systems are being built. In contrast to the traditional methods of handling rigidly timed video, networked media unshackles the system designer and allows for dazzling new possibilities of connectivity, storage, media management and media handling. This paper will review why the move to “IT”¹ based systems is so compelling for our industry. The paper will cover; what is networked media and its advantages; making IT work in real time; overview of SAN and NAS and a sample design for a large news production system.

Introduction

One of Sir Isaac Newton’s Laws states “*objects in motion remain in motion and objects at rest remain at rest*”. To an extent this has been the recent state of A/V system technology. The traditional methods of moving (SDI, composite,..) and storing video (tape, VTRs) assets are comfortable to the engineering staff, fit existing workflows and are proven to work. Some facility engineers feel, “If it’s not broken don’t fix it”. Ah, but another Law of Sir Isaac’s states that “*objects or systems will move in the direction of an applied force*”. So, what force is moving A/V systems today? The move towards an IT based infrastructure. But what are the advantages and tradeoffs? Let’s investigate these issues.

What is Networked Media?

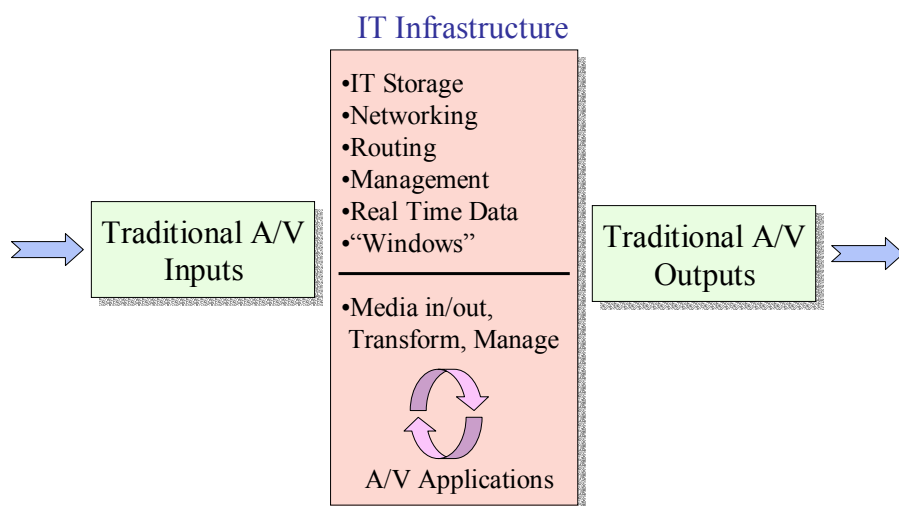
In contrast to traditional A/V equipment choices, networked media relies upon technology and components supplied by IT equipment vendors to move, store and manipulate A/V assets. With all respect to the stalwart SDI router, it is woefully lacking in terms of true networkability. Only by Herculean feats can SDI links be networked in similar ways to what Ethernet and IP (Internet Protocol)

¹ IT stands for Information Technology and its storage and networking concepts are universally used in business systems worldwide.

routing can offer. In summary the following demonstrates methods and concepts that use networked media;

- *Direct-to-disc* ingest, edit or playout
- 100% reliable File transfer methods
- *A/V-as-data* archive – not traditional video tape archive
- WAN, LAN connectivity and routing using Ethernet, Fiber Channel, SONET/SDH, commensurate switching and other IT methods
- Networkable A/V components – IT connected devices with optionally traditional A/V interfaces
 - Ingest ports, edit stations, data servers, caches, playout ports, proxy stations, A/V process stations and so on

Using IT in A/V Systems



The figure above is a high level view exhibiting IT technologies as the core of an integrated A/V platform. The philosophy is to ingest and output in standard A/V but do all other processing using IT as practical.

Networked Advantages

The Figure below illustrates six powerful reasons to embrace networked technologies.

The Networked Advantages

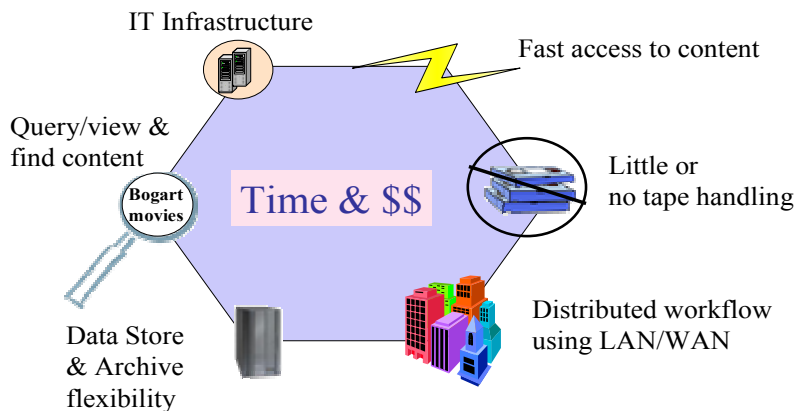


Figure 1

These advantages each save time and money per workflow compared to traditional methods. Let's review the elements clockwise starting at 1 O'Clock.

- Fast content access – Networked media enables local or remote clients to capture/playout/browse/edit A/V assets. Via direct-to-disc or file transfer, media is readily available for use.
- Little or no tape handling – Video assets are available on disc (on, near or offline storage) or via archived data optical discs or robotic tape access. Camcorder tape based acquisition will play an active roll for the next few years until disc/optical based camcorders are mature.
- Distributed workflow – This is a huge advantage. The old physical and geographical limits for a facility have been changed. With LAN and WAN connectivity, we can create a virtual facility. Imagine ingesting a clip in Tokyo, editing it in NYC and playing it out in London. This workflow is not a dream system. Systems exist that do similar workflows and more.
- On/off line disc storage of massive amounts of A/V assets will forever improve workflows. Additionally, by not archiving using a A/V format specific media type (D1 or D5 for example), media choices are expanded. Users may archive *video-as-data* and choose any number of format agnostic methods (DVD-ROM, data tape, long term offline disc and so on).
- Media Asset Management – Catalog assets, browse/query and put them to use. Much has been written about these systems and the IT transition will enable MAM solutions to bloom in many creative ways.
- IT infrastructure components – Ride the many performance and cost waves of this technology. (see below)

What are some of the benefits of using IT components in broadcast and post? There are three important curves that we will benefit from.

1. Ride the decreasing cost curve of Ethernet, Fibre Channel, switching and storage
 - a. \$30/port for basic LAN in 2003
 - b. IP Router cost/performance is ever improving
 - c. Hard Drive \$/GB approaching \$1 per GB for raw storage. 300 GB drives in 2003. One small array of 8 drives will hold 300 movies at 8 Mb/s. Storage density is one of the biggest drivers in the move to IT.
2. Ride the increasing performance curve
 - a. Gigabit Ethernet switching is a commodity item with 10 Gb/s Ethernet introduced in 2003. Fibre Channel is at 2 Gb/s today.
 - b. Moore's Law states that circuit density doubles every ~18 months for many types of electronic devices. This implies a bright future in terms of media processing horsepower.
3. Ride the expanding connectivity curve
 - a. Ubiquitous Ethernet and WAN connectivity. Local, metropolitan and international links connect facilities. DSL and Cable Modems at the low end and 40 Gb/s (SONET/SDH OC-768) at the high end.
 - b. Metcalf's Law predicts that the usefulness of a network of nodes is proportional to the square of the number of nodes. So, *the more connected nodes the better* is the mantra.

Let's take a tangent for a moment and investigate connectivity's very high end. Using Wavelength Division Multiplexing on optical fibre, researchers at Lucent Technologies/Bell Labs have proven that a WDM optical transceiver is capable of delivering ~40,000 Gb/s of data on one strand of Fibre. Using 1000 different wavelengths each carrying an OC-768 payload they postulate the astronomical rate of ~40 Tb/s is achievable per strand of fibre.

Let's assume that we have encoded an immense collection of MPEG movies and programs each at 4 Mb/s. At this rate, one could transmit 10 million different programs simultaneously on one single fiber. Since most fiber cables carry 200+ strands, one properly snaked cable could serve 2 billion homes each accessing a unique program. Ah, so many channels, so few people. Amazing? Yes, but tomorrow promises even greater bandwidths. What's the point of this hyperbolic illustration? Video distribution and production workflows will be greatly impacted by these major advances in connectivity. Fasten your seat belt and hold on for a wild ride.

Making IP Work in Real-Time

One criticism often raised against IT connectivity is its lack of true guaranteed isochronous delivery (like the common serial digital link, 259M). The cry is, "IT is not designed for video". Fair enough. It's not designed for video specifically but it may be coerced to behave in a civilized manner and meet even the strictest objectives of A/V system design. True there are some cases where SDI will be the preferred link.

But many common applications in use today can migrate to an IT infrastructure. What applications are amenable to using IT components? Figure 2 shows a segmentation of applications.

Applications That Are IT Friendly

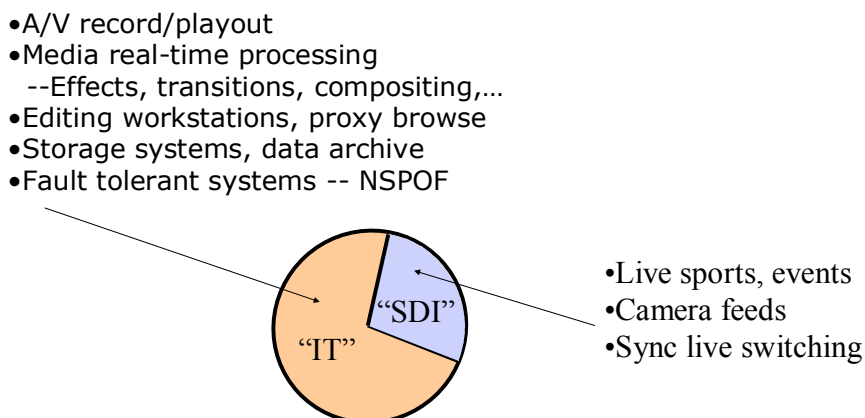


Figure 2

SDI will remain king in areas that require frame accurate switching between synchronized sources. It's possible to use Ethernet to carry camera feeds but the effort may not return sufficient rewards. For the immediate future SDI (and related links) has a solid place in live event production. But, many other applications can and do use IT components today.

So in the cases where IT components can be used, what special knowledge needs to be applied to a system design? If there is a 'trick' to persuading IT a component to meet our needs, it's proper use of data buffering. With the right amount and placement of buffering any A/V data jitters that are a direct result of LAN and WAN connections, disc latency, IP routers, data servers and other sources may be eliminated. Buffering is our friend on one hand and an enemy on the other. Buffering adds signal delay and in many A/V systems this delay must be accounted for and possibly hidden. So with judicious buffer placement, sizing and hiding strategies, IT components can be made to deliver and process A/V data in real time and error free. Of course it's the job of the equipment and system designer to manage the buffers so that the end user is not aware of their presence. Figure 3 shows the placement of the three most important buffers.

Let's call these buffers the **look behind buffer**, **look around buffer** and the **look ahead buffer**.

Making IT/IP Work in Real Time

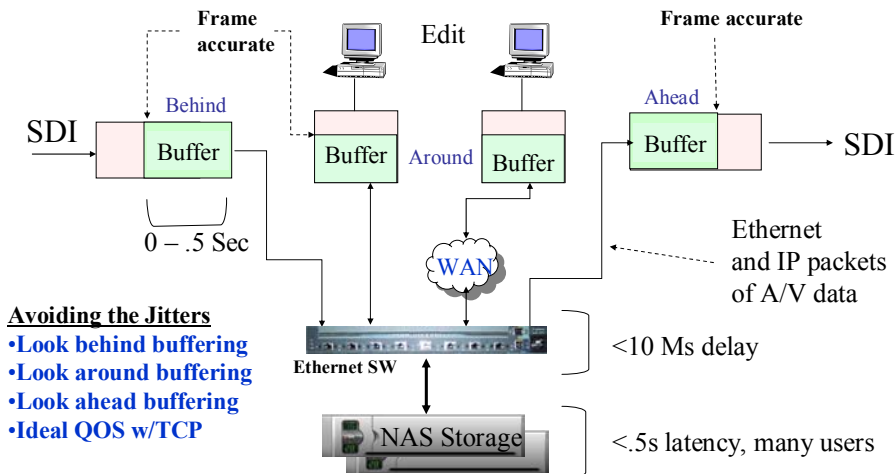


Figure 3

Let's consider the operational characteristics of each buffer. In all cases, a buffer should never overflow or underflow. The system in Fig 3 is a simple ingest to storage, edit off storage and play from storage architecture. It is a NAS based architecture but the reasoning applies equally to SAN based systems. NAS and SAN tradeoffs are discussed later in this paper. The **look behind buffer** stores any newly ingested A/V data that is to be written to disc. This buffer, as with the other two, is elastic in nature. One end fills at a constant rate (from the SDI video input) while the other end empties somewhat erratically as data is written to the storage array. As long as the buffer does not dry up or overflow, we can be certain that all input data will get reliably written to the storage system. The buffer contains data that is from the past or 'behind' the current input signal.

The **look around buffer** smoothes out any disc access irregularities during editing. Many non-linear editors support timeline jogging. Bouncing back and forth around a narrow range of frames is best accomplished from a local buffer. Imagine how sluggish the interface would feel if each jogged frame needed to come from a remote store. This is untenable and so the need for a local buffer. When one or more end stations are connected to storage via a WAN, the link jitter can become intolerable so proper use of buffer smoothing is of prime importance.

The signal playout device has a **look ahead buffer**. On queue this buffer precharges with a few seconds of A/V data then on command to play the buffer empties at a regular rate. This buffer operates in a manner similar to the **look behind buffer**. The queue command is analogous to a 'pre-roll' in a VTR. The buffer contains data that is waiting to be output, hence the **look ahead** moniker is appropriate. Also,

by using TCP/IP over Ethernet, the quality of data delivery is guaranteed to be perfect in a working system.

Under the normal workflow of ingest-edit-payout, Fig 3 is a practical system with outstanding performance. But if there is a need to ingest and payout immediately, the overall system throughput delay will limit some applications. So with A/V systems built on the backbone of IT, buffering is a necessary evil. Fortunately, the system-wide benefits far outweigh any deficits due to buffering.

For an additional level of performance and reliability guarantee, we have developed a buffer and bandwidth management protocol called the Media Access Server protocol. Each client node (ingest, edit, payout in Fig 3) connects to NAS storage using the MAS protocol. MAS encourages each client to be a good network citizen and not consume more than its allotted share of network or storage bandwidth. MAS also prioritizes which system processes can consume this bandwidth. For example, a node that is playing out real time A/V content has a guaranteed bandwidth assigned while a node that is doing a background file transfer is limited in how much bandwidth it can consume. Good network citizens make for reliable and predictable system performance.

Storage Access Methods

Storage systems make up a large share of what is considered IT infrastructure. They can be classed into two different categories. In each case network clients (PC, workstation, servers, dedicated HW and so on) accesses a remote storage system over some type of network technology. The most popular method of connecting to storage is called NAS for Network Attached Storage. The second method is called SAN for Storage Area Networking. Next let's examine their personalities and the pros and cons of each method. Our first examination will be of NAS.

Network Attached Storage

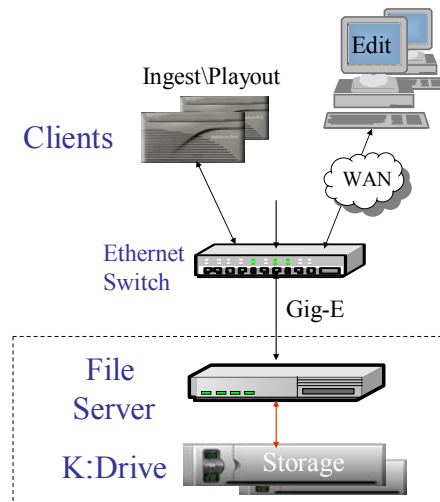


Figure 4

Networked Attached Storage – NAS

One of the foremost characteristics of a NAS environment is that storage may appear to any attached client as a networked drive, say the K: drive. The file server is also a critical element. Each client accesses the file server over an Ethernet/IP network and the server in turn accesses the actual storage arrays. The server contains the file system (Windows or Linux for example) that manages the storage array. It is this file system that all attached clients see in common. Of course security and access rights are available as required through the server's file system. When only one server is available, the server becomes the bottleneck in terms of data throughput. Normally, all the client data is funneled through an external Ethernet non-blocking switch.

A typical file server in 2003 will support 400-700 Mb/s of R/W data to the arrays. It is possible to create a cluster server (N smaller servers create one larger virtual server) that offers much more performance than that of a single server. These configurations are non trivial and may be expensive. Pinnacle supports a fault tolerant method of load balancing N servers between active clients. The aggregate throughput is in excess of 5 Gb/s.

In general there are two methods for NAS clients to access stored data. Let's call the first method **local file copy**. Fig 5 shows this case.

NAS File Serving and Editing

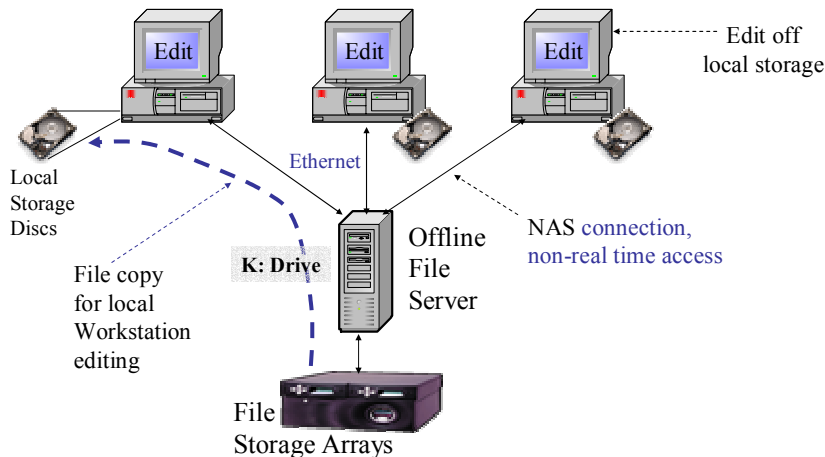


Figure 5

Consider an edit station that needs Clip_A that is stored on the K: drive. First the edit station will transfer all or part of the desired clip to the workstation's local memory. This operation may take place at slower or faster than real time. For example, a one minute DV clip encoded at 25 Mb/s will transfer to the edit station in 30 seconds if the file server supports a 50 Mb/s transfer rate. If there are interruptions or the transfer rate fluctuates, it's of little consequence as long as Clip_A eventually gets loaded into local memory. Once loaded, the editor can perform operations on the local version of Clip_A. Then, when the editing is done, the finished material may be file transferred to the K: drive.

This method has the advantage of allowing very loose network and storage array specs (data access latency, jitter, bandwidth,...). Since no clients need guaranteed and precise real time data delivery, network design, load balancing and disc access methods may be relaxed. So, the file delivery system is easier and cheaper to build. We have sacrificed real time A/V delivery performance with the consequence of slowing down the editing work flow.

A second NAS based method requires approximate real time performance of the arrays and delivery network. Let's call this technique **direct-to-disc** access. For this case clients access data directly on the arrays and edit as though the storage were local. No file transfer is needed as with the first case and the result is an improved workflow in terms of faster edit times. There are no free lunches and the quality of service for data delivery is strict and requires attention to system design and load balancing. Frankly, designing large systems (50+ nodes) requiring real time IT performance is part art and part science. Of notice is that such systems are being installed at A/V facilities worldwide on a regular basis.

Storage Attached Networking - SAN

A second method to attach to storage is called SAN. In this case storage appears as local attach to the clients. Figure 6 outlines a simple SAN architecture.

Storage Area Networking

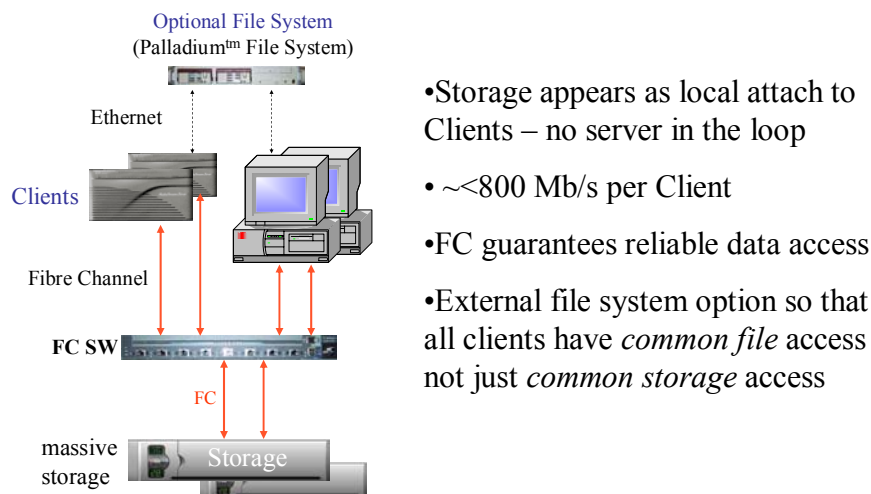


Figure 6

Key SAN Features Are;

1. Each client attaches directly to the storage arrays via Fibre Channel links. Each link can pass ~800 Mb/s of payload data although that does not imply that each client can consume that much data. The storage appears as local to each client.
2. The Fibre Channel protocol assures the reliability of the data on each link using hardware techniques. This is in contrast to Ethernet based connectivity where the TCP/IP layers are normally responsible for data reliability. TCP guarantees data reliability using software techniques and main CPU cycles are expended to this aim². So, Fibre Channel demands fewer client CPU cycles than NAS hence more processing power remains for client application software.
3. SAN provides clients with common storage access not common file access. This distinction is very important.

² There is a new trend for accelerating TCP software processing by moving this function from the main CPU to a secondary CPU on an Ethernet NIC card. This secondary CPU is responsible for managing the TCP/IP software tasks. 500 Mb/s Ethernet file transfers are achievable using this method with very little main CPU involvement. This performance will only improve.

Let's dig deeper into item #3. Since there is no file server layer as with the NAS case, individual clients cannot see the files written to storage by other clients. So, each client needs strict rules for accessing the storage else they will clobber each other's data. Pure SAN does not enable file sharing. So, to gain the benefits of file sharing between clients we need some type of Distributed File System that all clients can use in common.

We have developed the Palladium File System (PFS) to foster common file access on a SAN. It is Windows™ (other OS support too) compatible and is configured to appear as a J: drive to each client. Clients access the PFS controller via a low bandwidth Ethernet link. The PFS controller coordinates and manages all the file operations on the main A/V storage arrays. It administers all file read and write requests to the end that all clients share all files if desired. Importantly, no A/V data passes through the PFS controller. The PFS controller becomes the kingpin as all clients depend on it for file access. So the Palladium File System is designed to be fault tolerant using a variety of redundancy and synchronizing methods. There are many interesting techniques behind building a reliable, real time, high performance file system but they are beyond the scope of this paper.

Can a NAS system also use these file system techniques? Yes, when more than one file server is installed (see Fig 4) they may all access the FSC. In this way, each file server sees the same file system and as a consequence all clients see all files in common. A logical extension is to create a hybrid system where NAS and SAN clients peacefully coexist and share the same common storage and files. This is the basis of Pinnacle's Palladium Architecture. But why mix SAN and NAS? For flexibility in configuration, performance and reliability a hybrid mix is the best answer.

Regarding storage architectures, the Palladium Store (including the file system) was developed according to the principal of No Single Point of Failure (NSPOF). That is, no one component, link, internal element or software process may fail with a consequent loss of system functionality. We specify storage access operations under a very strict rule; full real time performance guarantee even when a single component is failing. This is a demanding constraint and one that requires world class design and test engineering. To implement NSPOF, the system must 'route around' failed components in real time to keep any video I/O glitch free. NSPOF may be achieved without a complete brute force mirror of all system components.

Our flagship product in this area is the MediaStream Server³ based on Palladium. This family of A/V servers can scale to 100+ channels with 11TB+ of storage. It supports MPEG media up to 50 Mb/s simultaneously in both SD and HD formats. MediaStream is a testament to mission critical playout servers. With over 1000 Palladium based broadcast channels on-air worldwide, MediaStream demonstrates the value of a scalable, bulletproof SAN architecture.

³ MediaStream is offered in two flavors; standalone (series 300, 700, 1600 I/O) and based on the common storage architecture of Palladium (series 900 I/O)

Sample System Founded on Principles of IT

Vortex News on Palladium

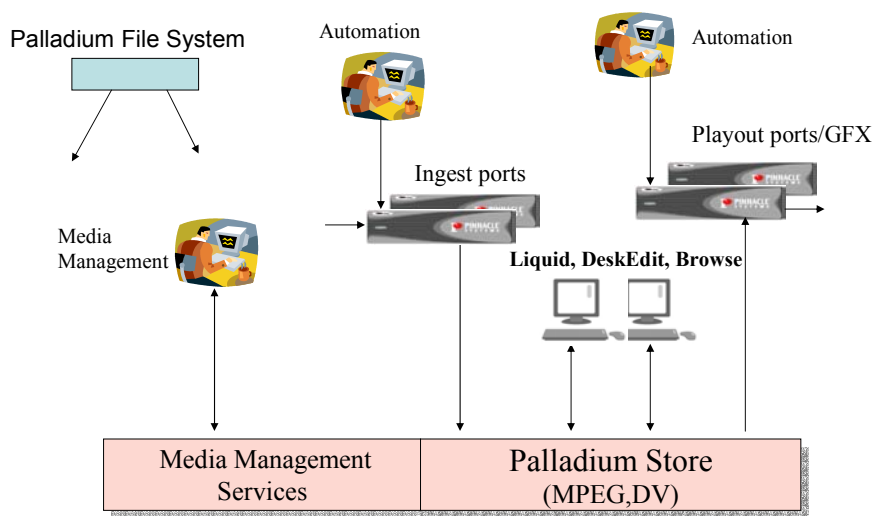


Figure 7

Figure 7 demonstrates a practical implementation of the Palladium architecture as applied to news story production. Some of the significant features of this configuration are;

- Common file system for all attached clients
- Hybrid SAN and NAS environment – clients may access storage via Ethernet or Fibre Channel links depending on design needs
- Fault tolerant (optional) common storage system
- Independent input/output modules
- News specific, Journalist edit stations (Vortex DeskEdit)
- Browse and storyboard stations using proxy clips
- Craft edit stations (Liquid blue, purple)
- Media Management applications and Metadata Registry. The Registry is open and 3rd party MAM vendors can interface to it.
- Integration with popular News Room Computer Systems (ENPS, iNEWS) and Automation solutions.
- Archive interface, system monitoring (SNMP) and a file transfer gateway (not shown)

Based on Fig 7, a typical workflow involves ingesting a clip into storage via an edit station or dedicated input port, editing the material on a variety of edit stations (News specific Vortex DeskEdit or a craft specific Liquid NLE), referencing the edited material in a rundown of a complete story and making the

edited material available for immediate on-air playback. All these operations run directly off the same common storage. No file transfers are required. This workflow is very efficient in terms of reducing the time to produce a story.

Of course there are other aspects of control APIs, metadata formats, file formats, file conversion gateways, compression support and so on. All these features of a system must be integrated with a combination of standard A/V and IT infrastructures. In addition, some solutions are more naturally formed by a combination of common storage and file import/export. Other models eschew the common store concept completely and depend almost exclusively on file transfer methods. There is no such thing as “the only correct way” to build an IT based A/V system. Expect to see many variations on the themes discussed in this paper.

Plus, there is new IT technology coming on the scene daily. Of interest is the new iSCSI protocol. iSCSI is designed to connect clients to storage arrays directly over IP networks. It is fast and efficient and marries Ethernet connectivity and the *SCSI command protocol* to create a new method for clients to access storage directly. It also eliminates the need for bulky SCSI cabling. iSCSI is just emerging so it will take a few years to see any meaningful penetration.

A real world example of this type of system is Time Warner’s 24 Hr/day news channel NY1 in NYC (and 7 other locations). It is based on Pinnacle’s Vortex News product. Journalists and editors produce NY1 news on a system that is similar to Figure 7. There is no video tape used at any step of the way except for the camera in-the-field acquisition stage. This style of system is revolutionizing news production and it will become the basis for a whole new generation of newsrooms.

Conclusion

We live in a time of the emerging world of IT based video systems. There is no doubt that IT components can be employed to meet the most demanding needs of many A/V production and distribution workflows. Due to legacy systems and signal interfacing requirements, IT will coexist with the traditional A/V infrastructure. A wide variety of products ranging from small I/O and edit appliances to 300 node systems are being built on the foundation of IT. Yes, IT is nudging its way into the ring of A/V system and product design. As it matures, it will become the core infrastructure on which the future of our industry will depend.