USING ISCSI IN VIDEO APPLICATIONS

BY: VICTOR KONSHIN & JEFF OKONCZAK

Application Engineers

ABSTRACT

This paper discusses the current video/audio workflow using direct attached storage (DAS) and how implementing a storage area network (SAN) addresses these limitations. Both Fibre Channel and iSCSI storage area networks are discussed with a focus on the advantages of using iSCSI to build a Digital Video SAN.

THE STORAGE AREA NETWORK – CONNECTING IT ALL

As video production infrastructure evolves to digital workflows, there is an increasing demand for storage, and a critical need to streamline the storage infrastructure. Traditional video production solutions involve multiple work stations with direct attached storage. This approach has serious deficiencies in scalability, utilization, availability and ease of management. To support today's digital production infrastructure there is a need to consolidate or pool the storage resources to enable shared access by multiple hosts.

In this networked environment, storage is consolidated and the work process of digitizing and editing content, enhancing projects with graphics and effects, color correcting and archiving content for future reuse, is greatly improved. Content is now centrally located, and the need to move large files across workstations with direct attached storage is eliminated. With a storage area network, multiple editors can work on a project simultaneously, reducing the completion time and simplifying overall management of a project. In addition, the total costs of ownership can be reduced while the availability of data is increased.

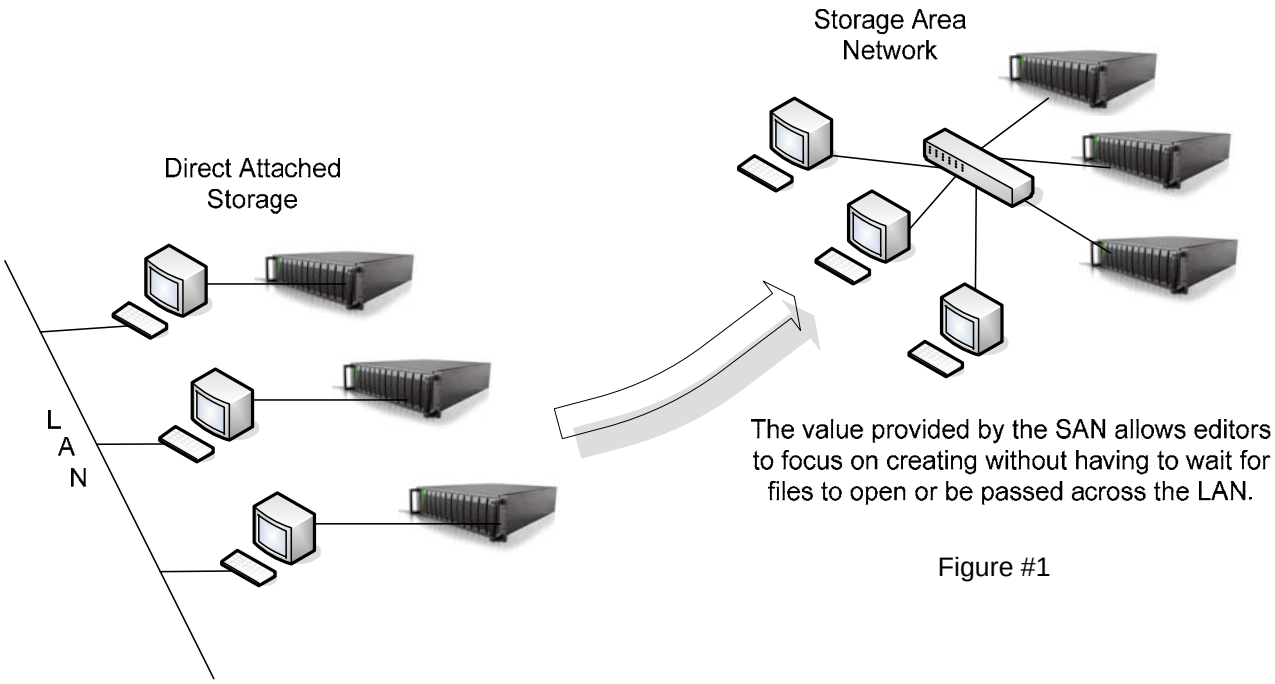
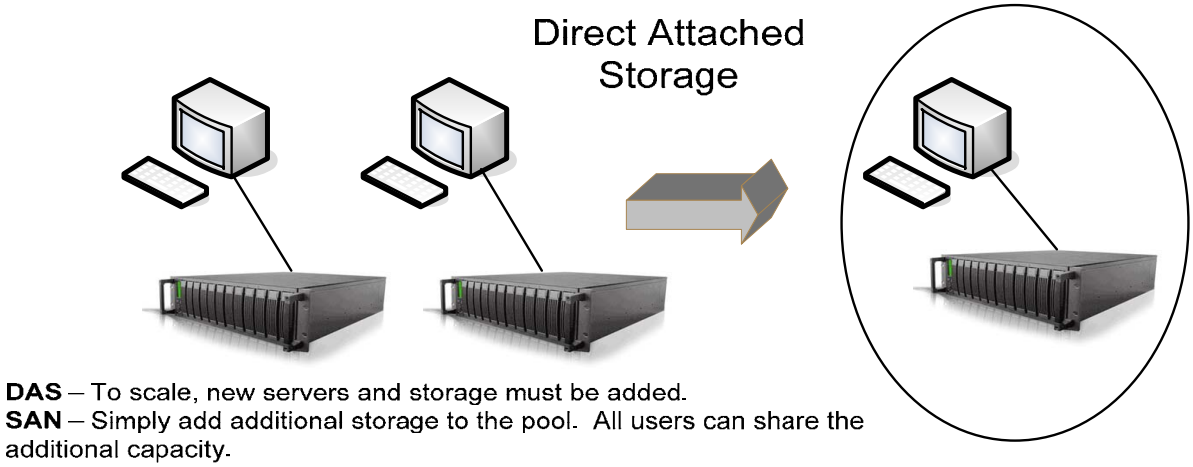


Figure #1

Imagine that your production storage environment was no longer limited by the number of devices that can be supported, the available performance/bandwidth or cable lengths. What if all the noisy local storage boxes were no longer inside editing suites? What if production managers did not have to add additional workstations to support additional storage needs? With current storage area network technology, the limits of direct attached SCSI are easily addressed. By building a SAN for a video environment it becomes simple to add new targets to the storage pool. Since the storage pool is shared, all workstations

on the SAN (not just the local workstation) can benefit from the increased capacity. Cable lengths can now be hundreds of meters allowing the storage to be centrally located outside of the suites.



DAS – To scale, new servers and storage must be added.
SAN – Simply add additional storage to the pool. All users can share the additional capacity.

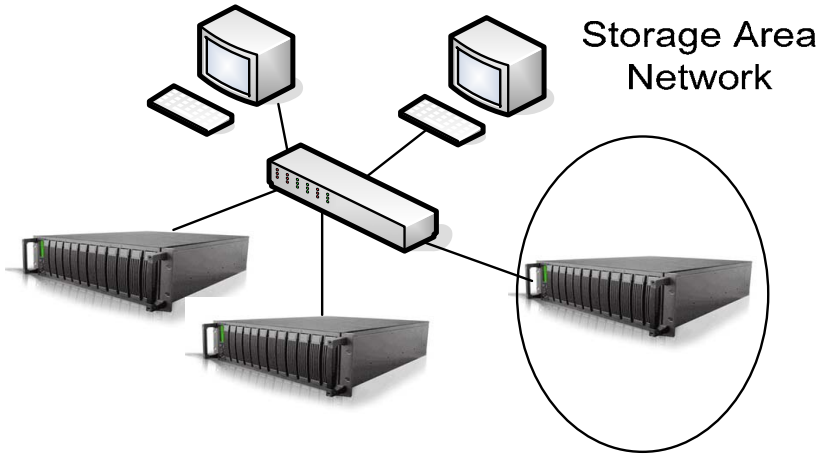


Figure #2

Storage utilization will also increase when migrating to a SAN. If one workstation #1 in a direct attached configuration has 95% of it's storage used but workstation #2 is using only 20% of its storage, there is no way for workstation #1 to take advantage of the unused space on workstation #2. In a direct attached scenario, additional dedicated storage would have to be purchased, installed and managed for workstation #1. With a SAN environment, the unused capacity on storage array #2 can simply be reassigned to any workstation on the SAN that requires more capacity.

By recouping capacity in existing resources and improving the collaboration between workstations, storage consolidation makes more efficient use of physical and financial resources, reducing the total cost of ownership.

Migrating to a storage area network can also result in reduced down time. When a workstation, host adapter, cable or storage device fails in a direct attached environment, many hours can be lost trouble-shooting and resolving the issue. In addition to affecting project deadlines, budgets are wasted as editors idly wait for the IT staff to resolve the issue.

A SAN environment provides an easy way to cluster workstations with multiple paths to the storage. Figure #3 shows an example of a configuration with multiple paths between the workstations and the storage.

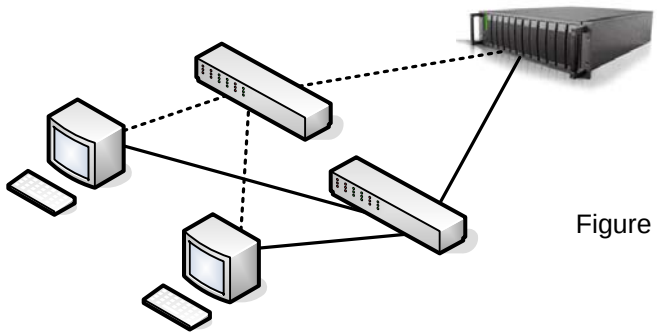


Figure #3

The key is that the storage networking protocols easily support multiple initiators connected to redundant ports on the same target device. Windows®, Linux and Macintosh® OS X all support multi-pathing and, in some cases, even load balancing. Connecting each path to a different switch, and then to independent ports on the storage array offers multiple paths to the same drive within the storage array. If a host adapter, switch, storage controller or cable fails, work can continue unimpeded through one of the redundant paths. If a workstation fails, the user can access the project from a spare station. This level of redundancy can scale for any size storage area network.

Deciding on the right storage technology can be a complex process. Many larger studios are consolidating their storage by using Fibre Channel technology to create their SAN. Fibre Channel works extremely well but the initial investment can prove to be costly for many video environments. Until recently, small and medium-sized production houses were often unable to deploy SANs due to complexity, price and cost. A better alternative may be IP storage using the iSCSI protocol.

WHAT IS ISCSI?

iSCSI is a storage protocol. Simply put, it is SCSI over Ethernet (TCP/IP). Similar to Fibre Channel, iSCSI uses efficient block level based SCSI commands and data. The difference is that iSCSI uses Ethernet as the transport. The additional control provided by the iSCSI protocol allows for multiple hosts to share the same storage targets. As discussed earlier, the ability to network storage and share it among multiple hosts allows for more efficient collaboration on projects that use large files such as video.

Instead of moving storage from computer to computer manually, iSCSI allows for inexpensive, high performance access to pooled storage.

WHAT ARE THE DIFFERENCES BETWEEN ISCSI AND FIBRE CHANNEL?

The main differences between iSCSI and Fibre Channel are performance and cost. With data throughput rates up to 400MB/sec., Fibre Channel is clearly a much higher performance technology. But that speed comes at a cost. The speed of a Gigabit Ethernet link (125 MB/sec.) adequately provides the bandwidth needed for most video formats. One exception is uncompressed high-definition video. Otherwise, iSCSI and Fibre Channel basically function in very similar ways allowing end users to consolidate their storage, easily provide multiple paths to their storage over long distances and share their storage across multiple hosts. Fibre Channel uses the Fibre Channel protocol to move SCSI commands and data whereas iSCSI uses the TCP/IP protocol to do the same.

One of the most significant benefits of using iSCSI is the cost savings that can be realized when compared to Fibre Channel. The cost of purchasing Fibre Channel host adapters can be replaced by Gigabit Ethernet controllers, which are built in to most workstations. In addition, highly reliable Gigabit Ethernet switches and cables are inexpensive compared to those required for Fibre Channel. Other benefits of using an Ethernet based protocol include flexibility, reliability, distance, ease of management and familiarity advantages associated with using the well-known Ethernet technology.

WHAT DOES AN ISCSI STORAGE AREA NETWORK LOOK LIKE?

An iSCSI SAN is implemented similar to a Fibre Channel SAN. Figure #4 illustrates multiple workstations connected to a Gigabit Ethernet switch using ordinary category 6 cables. Standard GbE network interface ports can be used in the workstations with the addition of an iSCSI Initiator. A software based iSCSI initiator is used to allow the host computers to communicate using the iSCSI protocol (block level transfers) over a standard Ethernet infrastructure.

The ATTO Xtend SAN iSCSI Initiator package for Macintosh® OS X consists of an iSCSI device driver, an initiator service manager, and a management interface. The iSCSI device driver is responsible for moving data from the storage stack over to the network stack, where it can be passed to and from the outside world using a standard Gigabit Ethernet network. The iSCSI initiator service manages all iSCSI ports in a host by aggregating device discovery information and security protocols. The management interface is designed to allow users the ability to configure specific protocol options and manage target devices through a GUI. There are third party iSCSI initiators available for the Windows®, Linux and Solaris operating systems as well.

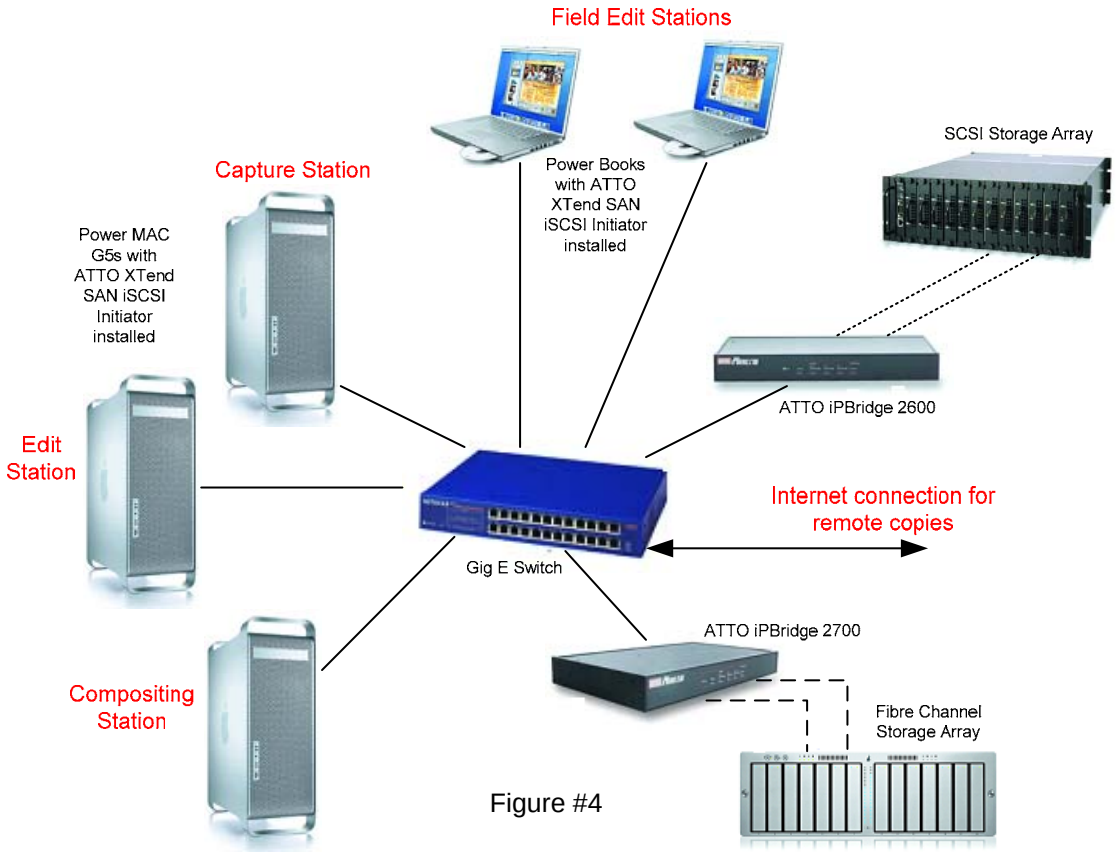


Figure #4

Once the hosts are connected to the switch, the storage is connected to the switch. One challenge is that native iSCSI storage devices no longer exist. The solution is to use an iSCSI-to-SCSI or iSCSI-to-Fibre Channel bridge in front of your SCSI or Fibre Channel-based storage device. The bridges offer much more than simple protocol conversion to the solution. They add value in the form of availability, connectivity, manageability, intelligence and management capabilities.

Bridge products allow for consolidated management of all drives using browser-based remote monitoring and management GUI's. These tools are typically supplied with every bridge and include features such as: SNMP management capabilities, advanced event and error logging with battery backed time stamp recording, link up/down monitoring, SCSI drive health monitoring, temperature and voltage monitoring and many more easy-to-use configuration capabilities.

Bridges also allow for in-field troubleshooting and diagnostics via a built-in trace log that provides bus analyzer-like capabilities. A key total cost-of-ownership factor to consider is how long it takes to identify

and resolve an issue once a hardware failure occurs. The troubleshooting assist capabilities provided in bridge products can reduce down time considerably.

Figure #5 provides a brief overview of ATTO Bridge products available.

The ATTO iPBridge™ 1550 provides one iSCSI to one Ultra320 SCSI port. The 1550 is capable of transferring data up to 100 MB/sec.

The ATTO iPBridge™ 2600 provides advanced diagnostic and management tools, better performance and four iSCSI ports with two Ultra320 SCSI ports.

The ATTO iPBridge™ 2700 provides four iSCSI ports with two 4-Gb Fibre Channel ports. It is capable of performance levels in excess of 360 MB/sec. It also includes advanced diagnostic and management tools.

Figure #5

The final element to consider, in the example of the video SAN, is the connection to the Internet. A router can be added to access the Internet or other network resources. This provides an easy way to extend your SAN long distances, while still using the block level transfers via iSCSI. Because iSCSI traffic between a host and storage can become quite heavy, especially for video applications, it is recommended that a separate Ethernet port is used to separate standard network traffic from storage network traffic.

ISCSI SAN CONSIDERATIONS:

Performance

When building SANs for video applications, drive latency is much more important than basic throughput. In the case of one host accessing one storage array, there is no significant issue with latency. However, when multiple hosts try to access data on the same physical drive, the drive's heads need to move back and forth between the different parts of the disk. With drive seek times in the milliseconds, this can compound very quickly causing performance degradation and dropped frames.

There are a couple of ways that this can be managed. First, is to have as many actual physical drives as possible. The more drives, the less likely it will be that more than one host will be accessing a drive at any given moment. Second, is to design the workflow to arrange your storage so that each user has their own space on a set of drives that only they will typically access. For example, the ATTO Diamond Storage Array has 24 drives. In a four user SAN, it might be ideal to have four-six drive RAID volumes giving each user their own set. If one user wants to access a file on another user's volume, they can copy the data from the other volume to their own. The fact that each user gets their own set of six drives maximizes performance while minimizing the chance that two hosts will access the same drive at the same time.

The second solution is to design your workflow to consider bandwidth to the storage array. For most video applications, 125MB/sec. of bandwidth between the workstation and the switch is sufficient. Bottlenecks tend to exist more often within the connection between the switch and the storage. If there is one link from the switch to the storage, the traffic from all hosts is squeezed through that one link.

VIDEO FORMAT	BANDWIDTH NEEDED
Standard definition: DV25	3.6 MB/sec.
Standard definition: DV50	7.2MB/sec.
Standard definition: 8-Bit Uncompressed	20MB/sec.
Standard definition: 10-Bit Uncompressed	30MB/sec.
High Definition: DVCPRO 720p	8MB/sec.
High Definition: DVCPRO 1080i	14.4MB/sec.
High Definition: 8-Bit Uncompressed 720p	120MB/sec.
High Definition: 8-Bit Uncompressed 1080i	150MB/sec.

To compensate, the ATTO iPBridge™ 2600 and 2700 both have four physical iSCSI ports providing an input throughput of up to 500 MB/sec. In addition, the multiple bridge ports provide for port fail-over protection.

One additional performance consideration is the TCP/IP packet size, known as maximum transmission unit (MTU) size. The standard MTU size for a TCP data packet is 1,514 Bytes. Larger sizes, called jumbo frames, are possible. Configuring the Network Interface Card for 9K or 16K jumbo frames will improve performance considerably by transferring more data per packet. This reduces the overhead by six times with 9K frames and 12 times with 16K. Not all GbE hardware supports jumbo frames so specifications should be carefully checked. It must be enabled on the NIC card in the host, the switch, routers and the target device.

VOLUME/FILE MANAGEMENT SOFTWARE:

Like Fibre Channel, iSCSI is a block level protocol. This means that the host talks directly to storage blocks on the device and is unaware of files, file systems or directory structures. Consequently, the host computer is unaware of any other computers that might be accessing the same storage. If one host writes a file to the storage, it's possible that another host can come along and overwrite it. Volume or data management software prevents this by coordinating where the various hosts write their files.

Data level management software administers on a file-by-file basis while volume management software protects on a volume level. The result is that data management software allows more than one user to write to a volume at the same time (can share volumes but not files); while volume level only allows one user write access to an entire volume at a time. In both cases many users can read from the same volume. Data management systems tend to be significantly more complicated as they usually require a custom file system and use a separate workstation to act as a controller. This controller uses permissions to coordinate where the various hosts write data to prevent corruption. This can add latency to a system. However, with file-level locking systems, there is little to no user intervention needed. They use their volume as they would any other. With volume level locking, the user must verify that they have write access before writing to the volume.

Apple's Xsan, ADIC's StorNext, Sanbolic's Melio and Rorke Data's ImageSAN are examples of file level locking products while CommandSoft's FibreJet and Charismac's FibreShare are examples of volume-level locking products.

SECURITY:

Because iSCSI uses TCP/IP, extra precautions need to be taken to prevent unauthorized data access. The easiest method is to not connect your iSCSI network to an outside network. This will make it impossible for outsiders to access the storage network. However, this also prevents other TCP/IP services from being used.

Other options to prevent unauthorized access include the use of Virtual Local Area Networks (VLANs), Access Control Lists (ACLs), Internet Storage Name Services (iSNS) or Challenge Handshake Authentication Protocol (CHAP). These are all standard networking tools that can be used to control which hosts have access to which target devices. An additional method to prevent access from the outside world is to block port 3680, the standard port used by iSCSI, on your routers.

iSCSI provides some additional options to increase data security. iSCSI digests will provide additional protection against data being corrupted in transit; however, because this environment is highly computationally intensive, there is a significant impact on performance and is not recommended for video or other high bandwidth application. Ethernet has built in link level error checking that makes digest unnecessary provided that Ethernet is used throughout the iSCSI network.

CONCLUSION:

iSCSI and Fibre Channel are becoming complementary technologies, each have a distinct place in the video environment. Fibre Channel provides for a relatively higher performance SAN while iSCSI is used where cost savings is a priority over performance. In either case, there are many distinct advantages over a direct attached SCSI environment.

iSCSI is a mature protocol that more and more video professionals are adopting the technology as a low-cost solution for networking their storage. There are a variety of products available – bridges, RAID arrays, tape libraries and initiators – for all the major operating systems. Working with the content creation market, ATTO Technology, Inc. has met the need for a low-cost iSCSI SAN solution through the introduction of iSCSI Initiator software for the Macintosh platform as well as bridges to connect Fibre Channel or SCSI storage to the iSCSI SAN.

These products provide users with the ability to tap the incredible benefits offered by iSCSI SANs. Users can take advantage of their existing Ethernet expertise and hardware to create a complete storage area network solution to consolidate, manage backup and protect their data, and network workstations directly to their storage.