DVS Whitepaper

SAN in Postproduction Environment: Requirements and Solutions

2006

1. Preface

In Digital Intermediate (DI) production SAN technology is on the move and with the worldwide implemented RoHS regulations, suppliers of SAN storage units will be changing to SAS (serial attached SCSI) or 4 Gb Fibre Channel (FC) by mid-2006. SAS has the double advantage with higher speed and lower price compared to 2 GB FC technology.

Taking these developments into consideration, this whitepaper investigates the new SAN technology. It will take you on an interesting journey from the history of postproduction to its predicted future. On this journey you will learn different SAN definitions, you will know why a SAN can solve challenging production problems and you might recognize, too, how to enhance the output of your facility. The day-to-day business requires flexibility, future-oriented thinking and innovative concepts. This whitepaper will help you in the decision making process and provides an informative grounding in general SAN technology as well as specific configuration setups and solutions.

Feel free to contact us in case of questions.

Kind regards,

The DVS Product Management

Contents

2. Postproduction Business: History and Future

Image 1 (Copyright: Pam Roth. Source: Wikimedia)

Traditional postproduction of films was ruled by chemicals and optical and mechanical processing. However, production of commercials have made use of video techniques for a long time. Editing was done by cutting the film and putting the pieces together manually, and video sequences were copied between multiple video tape recorders. Some special effects like wipes and dissolves could be achieved with special effects mixing devices which worked in real time. For this the original film had to be scanned and converted into an SD video signal at least once. Further processing was done with the video material. In most cases the final result was printed back to film.

This development was boosted by digital techniques in the 1970s. Digital video tape recorders (VTRs) and digital disk recorders (DDRs) replaced the analog VTRs due to the fact that disk recorders provided a faster access with almost no roll time and less wear. The serial digital interface, SDI, soon became the interface of choice. Today, HDTV is becoming the technology of every day business, being used in commercials and TV dramas, and even for the postproduction of some high-quality feature films. In addition to real-time HDTV equipment, digital computers and workstations are used for special effects work. These generally don't have a video interface but work with data files.

Image 2 (Copyright: DVS)

Video signal processing needs to work exactly at the speed of the video clock. If the task is too complex for a given processor, the task cannot be done at all. If the processor has more power than is required for the current task, the video clock will slow the processing down and the additional processing power cannot be used.

Future workflows will be based on data files where processing can be done at any speed. Complex tasks may be slower than real time, but simple tasks may be much faster. Processing speed is only limited by the power of the processing device which may be a specialized sequence processor or a standard workstation. Video equipment working in either SD or HDTV will be used less and less.

Key Facts: History
From celluloid to digital data, the history of film and video postproduction has undergone severe changes. Today's requirement for high data rates and data quantities imposes impressive challenges on the industry.

3. Definitions and Configurations

The postproduction of feature films and high quality commercials is typically done with uncompressed material. There is a tremendous amount of data when films are processed in HDTV or in cinema resolution with two or even four thousand pixels per line. Often multiple workstations need to work on the same project either to speed up processing or because different workstations are required for different software packages. Copying data between workstations via ordinary network structures is time-consuming. To avoid processing being slowed down by copy tasks, each workstation can be equipped with a local storage, to which large parts of the project are copied to, for example, over night. Now the workstations can access all frames quickly on the local storage.

This procedure has several disadvantages. The total storage of the individual workstations will be much bigger and much more expensive than a single storage holding the complete project. Copies of the same image may be needed on different workstations for different processing steps which may lead to coordination problems. If one operator has to wait for the result of another, then time for copying has to be added to the processing time.

Therefore a central storage, which can be accessed simultaneously by all workstations, may help to save time and money.

3.1 Central Storage with a Server-Client Configuration and NAS

The traditional solution for a centralized storage is a server-client configuration. The server is the only computer to which the storage drives are connected directly. All other workstations, the clients, are connected to the server via network. There is no direct connection between clients and the server storage. Of course, each client may still have its individual local storage directly attached. Nevertheless, every client can access the storage of the server. Hidden to the user, the communication is always between client and server, but to the user it looks as if he has direct access to the files on the central storage.

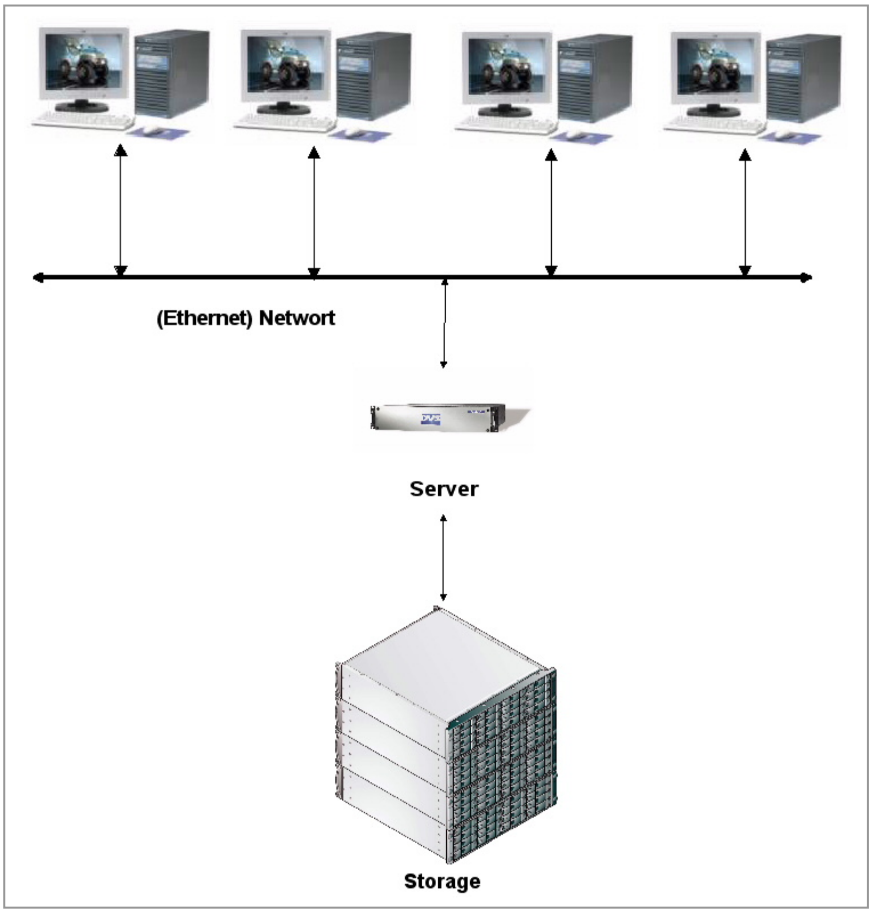

Image 3: Typical client-server configuration (Copyright: DVS)

Nowadays this configuration is also called NAS (Network Attached Storage) and often the storage and the server are mounted into a chassis not much larger than a normal desktop computer, making cabling and installation easy. While the standard network server also provides print services, naming services as well as others, a NAS system is dedicated to file services only.

In any case, the bandwidth of the server processor and the speed of the network connection can become a bottleneck, which makes a client-server/NAS configuration not suitable for real-time video and digital film applications.

3.2 Central Storage in a SAN Configuration

A so-called SAN (Storage Area Network) provides faster access to the data. All clients of a SAN are directly connected to the hard disk storage. The data transfer always takes place between clients and storage directly and there is no server that distributes the data which eliminates one bottleneck.

In practice the storage devices do not have enough connectors to connect many clients directly. Therefore so called switches are used to increase the connectivity. A metadata server helps to synchronize all clients. A more detailed description of both devices is given below.

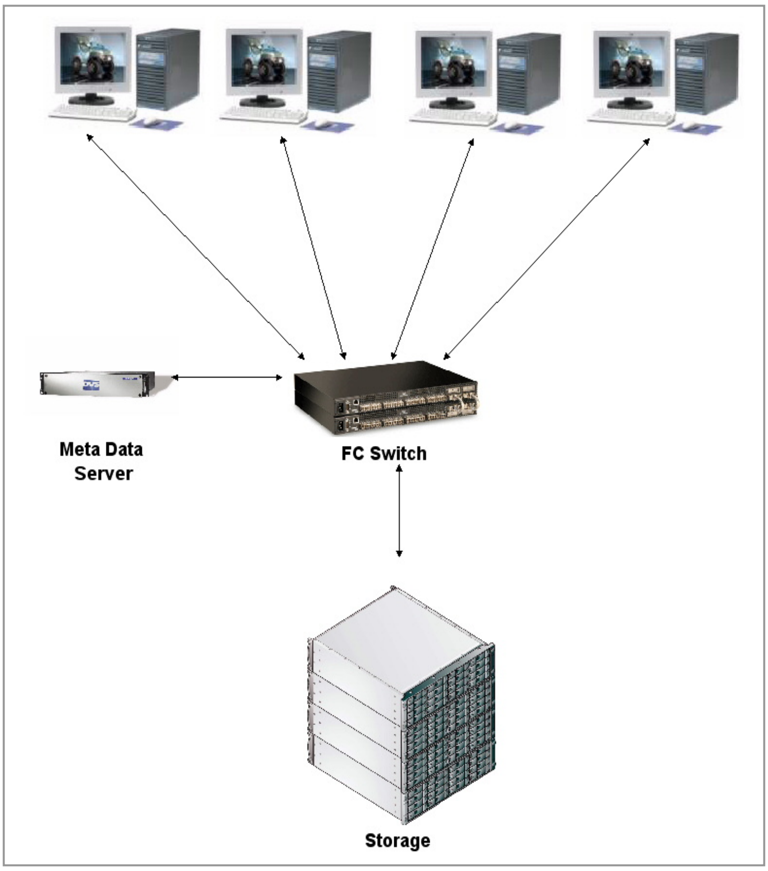

Image 4: Typical SAN configuration (Copyright: DVS)

3.3 Main Components of a SAN

A SAN normally consists of a combination of various components. The most important components are explained in the following text.

3.3.1 The Storage

The most important component, of course, is the storage itself which consists of a chassis with a RAID controller and a number of removable hard disk drives. Magnetic tapes, magneto-optical and optical devices are used for archiving purposes only, as their access time is too slow. Solid-state disks, though fast and reliable, are too expensive to be used for mass storages. There are two different types of hard drives.

Hard Drives for Home and Office Applications

These drives are optimized for high capacity and low price. Capacities go up to 500 GB and rotational speed is up to 7.200 rpm. The data rate of the media is between 30 and 65 MB/s with an access time of 8 to 9 ms. The number of reads between non-recoverable errors is about 100 times lower with it is as hard drives for servers.

The interface is either a parallel ATA (IDE) interface or its serial successor SATA and SATA-2. Many drives are available with either interface, but the technical data of the drives are identical.

Hard Drives for Servers

Hard drives for servers have to be much more reliable than hard drives for home and office applications, because they have to endure more accesses to the same data in everyday use. With these drives reliability, high data rate and low access times have the highest priority. The high data rates and low access times are achieved by higher rotational speeds. 10.000 and 15.000rpm are standard, 20.000rpm have been announced. Typically a manufacturer offers two rotational speeds and three different capacities for each speed, wherebye the capacities of the faster group have only half the capacity of the slower ones.

Overall, the higher reliability has to be paid for with lower capacities and higher price compared to the desktop drives. Whilst the capacity is only about half that of the desktop drives, the price per byte is about as twice as high. But the time between unrecoverable errors or the MTBF (mean time between failure) is about 100 times better. In addition they are recommended for 7/24 operating times, which is not always the case for desktop drives.

Depending on the manufacturer the drives are available with up to three different types of interface, namely the SCSI interface, the SAS interface and the FC interface. Again, the other drive specifications remain similar.

Overview of Hard Disk Interfacing Technologies						
	Home / Office			**Servers**		
	ATA 133	SATA	SATA-2	SCSI U320	SAS	FC
Max. Capacity [GB]	500	400	500	146/300	146/300	146/300
Rotational Speed [rpm]	7.200	7.200	7.200	15.000/10.025	15.000/10.025	15.000/10.025
Interface Data Rate [MB/s]	133	150	300	320	320	425
Media Data Rate [MB/s] (Value for 15K drive)	65-31*	61.4-29.8*	65-31*	142-85*	142-85*	142-85*
Media Data Rate [MB/s] (Value for 10K drive)	65-31*	61.4-29.8*	65-31*	89.3-46.8*	89.3-46.8*	89.3-46.8*
Access Time [ms]	8.5	8.5	8.5	3.7/4.7	3.7/4.7	3.7/4.7
Error Rate	1 in 10^{14}	1 in 10^{14}	1 in 10^{14}	1 in 10^{16}	1 in 10^{16}	1 in 10^{16}
GB = 10^9, MB = 10^6 * first block – last block.						

Image 5 (All data taken from data sheets from Hitachi Ltd. and Seagate Technologies)

3.3.2 RAID Controller

The RAID controller, where RAID stands for Redundant Array of Independent Disks, combines the hard disks of a storage unit. To the computer they look like one large drive. But the RAID controller also stores additional information, so-called parity information, on the drives. This makes it possible to recover the data of any disk if it fails. Of course, the information needs space and reduces the net capacity. Combining multiple hard disks could improve the speed, but the calculation of the parity also takes time. Therefore, most RAID controllers are slower than a simple stripe set as provided by many file systems.

If a drive fails, the original data can be recalculated using the parity information. With many RAID controllers the access time is slower than in normal operation when all drives are operational.

If a drive failure has been detected, the defective drive needs to be replaced. This can be done by swapping the drive manually or by switching automatically to a so-called hot spare drive seated amongst the other drives and unused so far. Then the recovery run will start and recover the data of the lost drive, writing it back to the new one. As described above, during recovery the data can still be read but at a reduced speed.

3.3.3 Metadata Server

A SAN does not need a server that transfers the data to the clients. But it needs a metadata server which makes sure that all clients use the same directory information. Every time a client creates or deletes files, all other clients need to be informed about this event.

The metadata server is a software package that can either run on a dedicated computer or on one of the clients. For real-time applications and if high reliability is required, the metadata server should run on a dedicated computer. This computer then is the metadata server.

As the metadata server does not actually move any data, it is less likely to become a bottleneck. But in real-time applications the reaction time will be critical.

The metadata that is handled by the metadata server should not be confused with the metadata that is used in video and film applications. All additional information to film and sound material is called metadata as well. It may include any information collected during shooting or postproduction. Some of the most important metadata are timecode and keycode, which provide a unique time stamp to each captured image and help to match video with audio. This type of metadata is treated like all other data by the metadata server.

3.3.4 Interfacing – Fibre Channel

Fibre Channel is the de-facto standard for the connection between clients and storage. It is a point-to-point connection that transmits via the SCSI protocol. Though the name implies an optical interface there are also copper versions available. For connections between disk units normally an optical interface is used. Inside a chassis the copper version is used, for example, Fibre Channel hard drives have a copper Fibre Channel interface.

There are interfaces for 1, 2 and 4 Gbps speed. Interfaces of all speeds can be mixed together. The actual transfer speed will be the speed of the slowest interface.

3.3.5 Fibre Channel Switches

As Fibre Channel is a point-to-point connection, switches are required to configure complex networks where one storage is connected to multiple clients. The most common switches have 8 to 32 ports. Some of them have expansion ports to combine two or more of them to operate as a single switch with multiple ports.

For each port the routing can be configured individually defining the ports that can be accessed from another one. This way it is possible that some clients can communicate with all storages to the switch while other clients can access only one storage.

The data rate of a switch may be limited and may be below the total data rate of all ports. Then, a switch may become a bottleneck. But it is easier to increase the data rate of a switch by combining switches via expansion ports than to increase the data rate of a file server; typically the bottleneck in client-server configurations.

3.3.6 SAN File System

'Normal' file systems have no integrated metadata server function. Therefore, either an extension to the file system is required or a special SAN file system.
An extension to the file system adds to the communication between the clients but leaves the control of how the data is stored on the hard drives to the native file system.
Typically no dedicated metadata server hardware is required, but one client has a kind of master function. DVS tests have shown that the master client has the highest priority and gets the best performance, which is a problem if more than one real-time stream is required. Another disadvantage may be that file system extensions are limited to the operating system of the native file system. Only clients with the same operating system can use them.

If clients with different operating systems need to be integrated to a SAN, a dedicated SAN file system may be a better choice. The SAN file system is a complete file system on its own, but it can be accessed from the clients in the same way as their native one. The most popular SAN file systems for video applications are ADIC's SNFS (Stor Next File System) and SGI's CXFS (Clustered Extended File System). Both are available for clients with various operating systems and both usually require dedicated metadata server hardware.

Some other SAN file systems and file system extensions on the market are detailed in the following table (in alphabetical order):

StorNext	from ADIC
Xsan	from Apple
SAN File System	from DataPlow Inc.
TotalStorage SAN File System	from IBM
Lustre	from Linux
GlobalFileSystem	from Linux Red Hat
ImageSAN²	from Rorke Data Inc.
Melio	from Sanbolic
CXFS	from SGI
QFS for Solaris	from Sun
SANergy	from Tivoly

Image 6 (Copyright: DVS)

3.3.7 Miscellaneous

The communication between the clients and the metadata server takes place via a normal network. For optimum performance there should be a dedicated network for this purpose. An additional network for maintenance and diagnostic may be suitable. These networks will require network switches and network cables as any other standard computer network.

Fail-over configurations for the metadata server need remote controlled power switches.

All these components must be provided with space in the rack and should not be forgotten when the rack space is planned.

Key Facts: What is a SAN?
A SAN is a Storage Area Network. Several setups are possible – from a classic server-client configuration to a Fibre Channel setup. Depending on the individual requirements in a facility, different configuration solutions have to be taken into account.
The DVS-SAN offers a wide range of possibilities. Both its design and its software setup enable the facility to handle large amounts of data with flexibility on a rock-solid platform.

4. Special Requirements for a real-time SAN

If the technical director of a post house is considering using a SAN, he normally knows the required storage capacity and probably how many workstations need to access the data. But buying a SAN for real-time video or digital film applications is not a simple question of "How many gigabytes do I get for my money?"

4.1 How much Storage do I need?

The calculation of storage capacity seems to be easy but has some pitfalls and is worth consideration.

The first thing to know is the resolution of the material. It is obvious that HD files are larger than SD files, and film resolution with two thousand or four thousand pixels per line take even more space. The file format has an enormous influence on the size of the files. Of course, compressed formats require less space than uncompressed formats, but color coding and quantization also affect the size.

Film projects typically will be stored in RGB 4:4:4. Video material in general is YUV 4:2:2. Nevertheless, it may be necessary to use an RGB storage format for video projects. Maybe the software used does not support YUV 4:2:2 or some processing steps require an RGB color space. Depending on the file format, RGB 4:4:4 requires about 50% more storage than YUV 4:2:2.

Quantization is another factor that has to be considered. Some file formats or some programs always use a multiple of 8 bits per color component rather than a compressed format even if the file format supports packed storage. TIFF files for example normally use 8 or 16 bits per color though the format specification allows for 10 bits. This requires about 50% more capacity than using DPX, which uses only 32 bits to store a 10 bit RGB pixel.

File Sizes and Storage Requirements			
Formats	Net size per image	Typical Storage requirement for one hour of video or digital film	
	[10^6 bytes]	Net size [10^9 bytes]	DPX File Size [10^9 bytes]
SD, 720x486, 30 fps, YUV 4:2:2, 10 bit	0.8748	94.45	101.74
HD, 1920x1080, 30 fps, YUV 4:2:2, 10 bit	5.1840	559.87	598.08
HD, 1920x1080, 30 fps, RGB 4:4:4, 10 bit	7.7760	839.81	896.68
Film, 2048x1080, 24 fps, RGB 4:4:4, 10 bit	8.2944	716.64	765.12
Film, 2048x1556, 24 fps, RGB 4:4:4, 10 bit	11.9501	1032.49	1102.03
Film, 4096x3112, 24 fps, RGB 4:4:4, 10 bit	47.8003	4129.95	4405.28

Image 7: File sizes and Storage Requirements. (Copyright: DVS)

DVS-SAN, of course, can handle all file formats, while DVS video applications like CLIPSTER®, Pronto2K.2 and ProntoHD.2 support a variety of them in real time, including compressed and uncompressed RGB and YUV formats.

Sometimes it may be necessary to have old versions of the project available as the customer may change his mind and may want to revive an old version that was discarded some time ago. So additional storage should be reserved.

The parity of a RAID configuration and hot spare drives, which are all available for DVS-SAN, increase safety and reliability but reduce the net capacity and should be taken into consideration if this kind of data protection is required. RAID 1 would need twice the net capacity. RAID 3, 4 or 5 need one additional drive per RAID pack. Mostly the size of the RAID packs can be configured. Large RAID packs need less overhead, but makes it more likely that another drive fails before a recovery is completed. Of course, the RAID pack size can be configured for the DVS-SAN. Normally the size is chosen as a compromise between overhead and reliability.

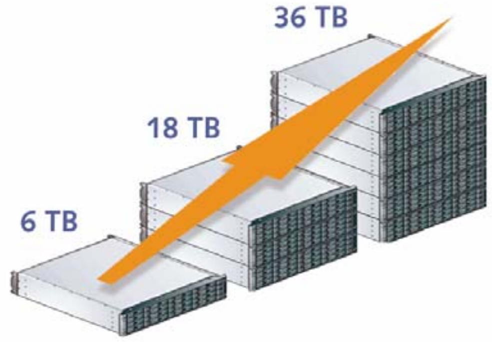

Image 8: DVS-SAN: grows with your requirements (Copyright: DVS)

But even the best planned estimation of storage capacity changes as the workflow changes or even better, if the business grows. To cater for the DVS-SAN can be expanded easily at any time. In most cases new storage can be integrated without having to delete and reformat the existing storage.

4.2 How many Clients shall be connected to the SAN and who needs real-time Access?

This question is not too difficult to answer, but has an impact on costs. Every client that is connected directly to a SAN needs a SAN license file system and at least one port on the FC switch. Clients needing higher data rates for HD, 2K, 4K film resolutions may require additional ports, as two or more links may be necessary for a single data stream.

If a large number of clients need non-real-time access, a file server may be a cost-effective solution. The file server is a client of the SAN. It is connected directly to the storage and needs a SAN file system license, too. But all clients that are connected to the file server neither need a SAN license file system nor a port on the FC switch. Both can help to reduce costs.

DVS-SAN optionally offers file servers to keep costs down in applications where many workstations or render nodes need access to the data simultaneously but not in real time.

4.3 Which Data Rate does my SAN need?

This question is the most difficult to answer. Probably there will be data that needs real-time access. But there may also be data where speed is not an issue. If both types of data are stored on the same storage, the non real-time access may interfere with real-time transfers. To avoid dropped frames the storage should be over-specified in terms of data rate. As a fast storage for real-time access is more expensive than a slow storage, it is a good idea to split the storage into real-time and non real-time or near-line storages. The non real-time data does not occupy the expensive real-time storage and costs for over-specifying the data rate are avoided.
This can reduce costs substantially.

Visit www.dvs.de and check the data rate calculator (see 'Tips & Tricks').

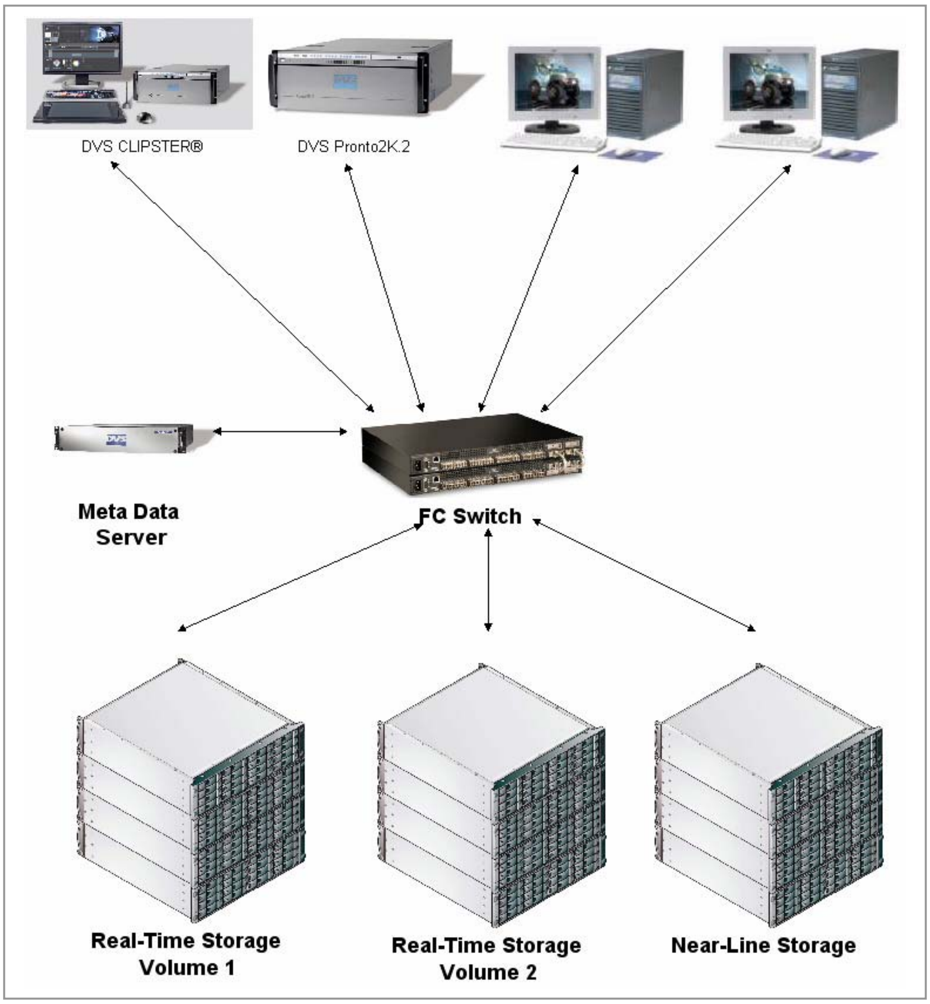

Image 9: DVS-SAN configuration with multiple real-time volumes and a separated near-line storage (Copyright: DVS)

Typical Data Rates		
Format	Net* Data Rate [10^6 Bytes]	Gross* Data Rate [10^6 Bytes]
SD, 720x486, 30 fps, YUV 4:2:2, 10 bit	26	28
HD, 1920x1080, 30 fps, YUV 4:2:2, 10 bit	156	166
HD, 1920x1080, 30 fps, RGB 4:4:4, 10 bit	233	249
Film, 2048x1080, 24 fps, RGB 4:4:4, 10 bit	199	212
Film, 2048x1556, 24 fps, RGB 4:4:4, 10 bit	287	306
Film, 4096x3112, 24 fps, RGB 4:4:4, 10 bit	1,147	1,224

* Gross data rate is calculated for a storage format where 30 video bits are occupying 32 bit, as most file formats like DPX, TIFF and others do. Net data rate is assuming there is no overhead in the storage format.

Image 10: Typical data rates (Copyright: DVS)

4.3.1 Peak and sustained Data Rate

Peak data rate is the data rate that can be achieved for a very short time only but not continuously. Sometimes it is defined as the data rate that will never be outperformed. Data sheets often specify peak data rates.

Sustained data rate is the average data rate over a longer period. Mostly it is not defined exactly how long. So it is not guaranteed that the actual data rate does not drop below the sustained data rate for a short time. In real-time applications this may cause problems. It cannot be recommended to use a standard benchmark application in order to measure the data rate of a storage. It is rather useful to run the storage with the intended software application which will be used in the future and then to measure the data rate of the running system. Then, the data rate given will be exact.

4.3.2 Real-time Storage

The necessary data rate of the real-time storage of a SAN is mainly defined by the data rate of the real-time clients. A first guess for the data rate requirement would be to sum up the data rates of all real-time clients and use a benchmark program to select a SAN that is fast enough.

Unfortunately in reality this may not give a reliable result. Reading or writing a single stream in large blocks will result in a much higher data rate than reading or writing small blocks in random order. Some applications use file formats like QuickTime or AVI that store a complete sequence including several audio channels in one file. If the file is not fragmented this comes close to the benchmark with sequential read or write in large blocks. But if two or more independent reads or writes take place at the same time, the total possible data rate usually drops. One data stream with 250MB/s is easier to achieve than five independent streams with 50 MB/s each.

Things may get more complicated if the video sequences are stored in a large file sequences, each containing just one frame, as in a typical DI workflow. If the files are stored continuously in one large block on the drives, the possible data rate is about the same as in the case of a sequence that is stored in one file. But if the files are scattered all over the hard drives, the data rate may drop dramatically. Scattering or fragmentation typically takes place during postproduction when single files are modified for retouching or when the sequence of the images is changed during editing.

While standard de-fragmentation tools can easily make sure that each file lays more or less continuously on the drive, there is no standard tool that takes care about scattered file sequences where every frame is at a different location on the hard drives.

Unfortunately many tools used in the postproduction of digital films use single frame file formats. It makes handling of large sequences much easier and programs that have been developed for still pictures can be used on single frames or sequences without file conversion.

To avoid problems caused by scattered files, DVS has developed a tool that comes with DVS-SAN to "de-scatter" file sequences. Moreover, the innovative data manager by DVS, Spycer™, is an interesting solution in the defragmentation process. See chapter 6.2 Spycer™.

Audio also may reduce the usable data rate. The 750kB/s (2 x 24 bit/sample, 48 kHz) of a stereo audio stream is very low compared to the several MB/s of video or digital film, but the additional seeks will need waiting time when no data can be transferred to or from the hard drives. Some real-time applications like CLIPSTER® and Pronto2K.2/ProntoHD.2 allow one independent file for each audio track and 16 or more tracks can be used for a single video timeline. The greater flexibility has to be paid for with a higher data rate requirement for the storage.

Video applications such as CLIPSTER® that provide real-time transitions even in 2K resolution will double the data during the transition.

Sometimes 'unexpected' data rates may occur which can jeopardize a real-time transfer. 'Unexpected' data rate can be caused by a scheduled copy process running in the background from time to time, which the operator of a real-time transfer is not aware of. To detect data rate overflows due to 'unexpected' transfers, DVS-SAN provides automatic e-mail notification if the data rate of all current transfers gets close to the maximum for which the storage is designed for.

Besides pure data rate, latency is another important point that influences the real-time performance of a SAN. The access time of a certain block of data is not exactly predictable as hard drives include mechanical parts. Depending on whether a seek is necessary and where on the next track the requested data is located, additional time consuming revolutions of the disk may be necessary. Retries to recover read or write errors may add additional delays. Professional real-time video systems will buffer some of these unpredictable delays. But larger buffers will result in slower response times to user commands. Remote controllers for VTRs and DDRs sometimes require a limited response time which reduces the maximum buffer size of a DDR that reads its data from a SAN.

So even if the average data rate is high enough, it may happen that the real-time transfer is disturbed by occasional high latencies. Tests at DVS have shown that SATA drives, which commonly have higher capacities per disk surface and therefore need more accurate tracking, suffer more from occasional higher latencies than FC or SAS drives.

Some well-designed video systems like CLIPSTER® and Pronto2K.2/ProntoHD.2 will use dynamic buffer sizes to optimize user-friendly response times with robustness against latency of the drives.

To minimize latency of drives and to give the highest reliability DVS-SAN always uses FC or SAS drives for real-time storage applications. SATA drives are used for near-line storages only where fast access is desired, but here occasional latencies do no harm.

When configuring a SAN, it may be wise to have two or more volumes on independent storages that can handle one real-time stream rather than a single storage that can handle two or more streams in real time (see image 9). Independent storages make sure that there is no interference between the transfers, providing additional safety against accidental data rate overflows. The disadvantages are that the storage is split into two or more segments and the user has to decide which storage should be used for which project.

DVS-SAN provides storage capacities and data rates in a granularity that makes it easy to configure. Configurations can either provide a single volume with high data rate to handle multiple 2K streams including audio in real time, or independent volumes, each with just one high resolution stream, without too much overhead in data rate and costs.

Image 11: DVS-SAN (Copyright: DVS)

4.3.3 Near-line Storage

Per definition the data rate of the near-line storage is less important. But as it is mostly used for archiving no compromise should be made regarding data protection. The consistent low latency of FC or SAS drives is not necessary. SATA/SATA-2 drives that offer a much better price per byte are a suitable choice, but RAID protection should not be omitted.

Of course, there are near-line storages of DVS-SAN available at an attractive price, providing always the same reliable RAID technology and remote diagnostics as the real-time storage.

4.4 How do I organize my Data?

A SAN normally provides a lot of storage capacity that can hold a huge amount of data. This may lead to the problem of how to find data. A first approach will be to store clips into folders named by project and subfolders with the names of the clips or a processing step. But sometimes the name of the clip is not known; the user just remembers the content. Or a user needs a clip of a particular timecode or keycode or is looking for a file type. In all these cases tools are helpful that can search for various meta information and show a thumbnail preview of a clip.

Also moving data with standard copy programs is not always efficient.

DVS-SAN comes with the powerful Spycer™ that does all the above plus many other useful functions. See chapter 6 for more information.

4.5 What operating Systems can my Clients use?

A typical postproduction house uses many different tools that run on different workstations. Normally not all of them use the same operating system. So the SAN should support all operating systems that are used by the customer.

DVS-SAN can be connected to all common operating systems including Windows® XP, Windows® Server 2003, 32 and 64 bit Linux®, IRIX® and other UNIX versions, and MacOS. Clients with all operating systems can be connected at the same time and access the same data.

4.6 How to handle Failure

Even the most reliable system may fail at some time. The most common is a drive failure. Therefore, DVS-SAN uses RAID 5 protection by default. If a drive fails, an alarm will sound but the data will still be available for processing. Real-time performance may be reduced though. Hot spare drives make sure that the recovery run can be started immediately after the drive failure was detected with no immediate need to replace the broken drive. Once the recovery run has finished, the system will be back to full performance and the failed drive can be replaced at a later time.

On request DVS can provide configurations that have no loss of performance even in the case of a drive failure. This is done by using RAID controllers that always omit the slowest drive of a RAID pack and always work in data correction mode. This way it does not matter if one drive is just a little slower or does not work at all. The performance will always stay at a maximum.

As mentioned above, FC and SAS hard disks have a lower error rate than SATA drives. So using a SAS or FC configuration is another way to improve reliability. DVS-SAN is usually configured with FC or SAS drives for real-time applications and with SATA drives for near-line and archiving applications.

Additional safety can be achieved by using dual controller configurations and a redundant metadata server with fail-over capability. If one controller or one metadata server fails, another one will take over automatically. Of course, DVS-SAN can provide both.

Fibre switches normally are backed up in a cold spare configuration. If a switch fails, only few cables need to be reconnected and the DVS-SAN will work as before.

4.7 What Support can I expect if my SAN fails?

If there is a problem, fast diagnostics and problem solving is essential. The diagnostics can be a major problem in a real-time SAN environment with attached video systems from different manufacturers. In general, storage manufacturers don't have much experience with real-time video systems and video manufacturers don't always have a lot of experience with SAN storage configurations. In case of a problem, it may happen that storage and video manufacturers blame each other! DVS is a manufacturer who has many years of experience in video **and** storage equipment manufacturing and who understands the workflows and requirements of postproduction. Remote diagnostic over Ethernet and automatic e-mail notification by the DVS-SAN help to keep down-times to a minimum.

Key Facts: Why do I need a DVS-SAN?
There are several demands on a real-time SAN expects several requirements. Of course, it is essential to know how much storage is necessary. But there is more to it than the famous money versus storage calculation. There are more questions to be asked and answered: How many clients will have access to the SAN? Is the SAN needed with a real-time storage or a near-line storage? What is the best way to organize the data? Are there different operating systems in use? Is there any protection against failure – and what support is available when needed? Only if these questions are considered, can a suitable SAN solution be implemented that meets your requirements and enhances your workflow.

5. Application Examples

5.1 A virtual Telecine

A SAN is a very flexible unit. There are almost endless possible applications to fit any requirement.
Below is shown a relatively simple setup with just one real-time film or video stream. This application example is a virtual telecine configuration. The SAN must deliver a sustained data rate even when all other clients are accessing it.

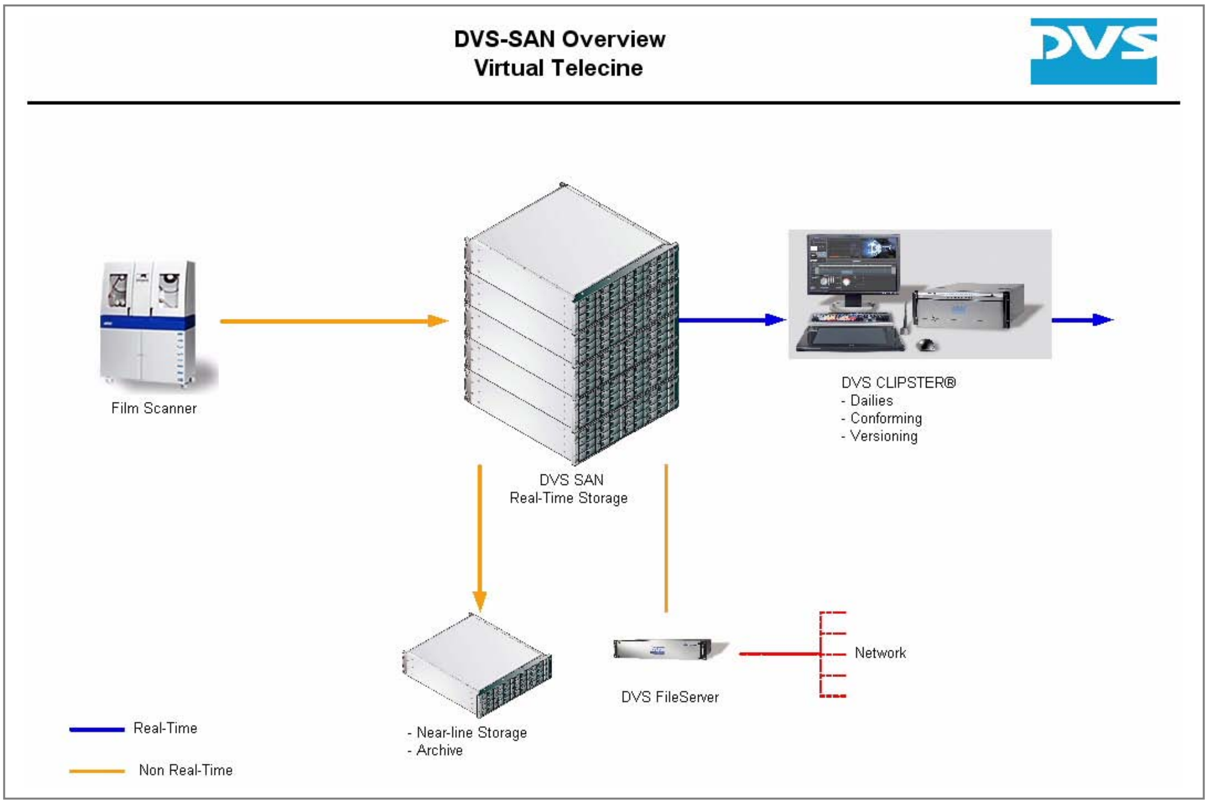

Image 12: A SAN configured as a virtual telecine (Copyright: DVS)

The film is transferred by a film scanner to the central storage via a data interface in non real time.
While being transferred from the scanner, the images can already be used by other devices. Here for example, a CLIPSTER® which is doing the conforming of the transferred reels. After the conforming the material can be played out in real time for inspection or color grading.

Also shown here is an archive that could be a near-line storage, a tape library or a VTR.
In parallel a FileServer can make the images available for authorized network clients which could be for example, the control station of the film director or DRS workstations.

5.2 Real-time SAN and near-line SAN

A significantly bigger application is shown in this example. Here the DVS real-time and near-line SANs are the center points for several departments of a facility.
The media is collected either by a real-time ingest from VTRs via an ingest station like e.g. ProntoHD.2 or as data from a film scanner or a telecine. Full resolution and proxy versions of the data can be stored on the real-time or near-line SAN as well. The proxy will be used for the offline editing afterwards where an EDL is created.
Once the data is transferred to the SAN and the EDL from the offline editing is available, the conforming and finishing devices can directly access the full resolution data, concurrently in real time.

Applications with lower speed requirements like render farms or digital restoration systems can access the data in parallel. Also the parallel access of multiple network clients via a DVS FileServer is available. The requirements for the real-time SAN in this case would be multiple real-time streams of HDTV or 2K film in parallel for the ingest, conforming and finishing, while having multiple non-real-time access to the same storage space. The only copy processes needed would be the transfer of the data from the render stations to the SAN.

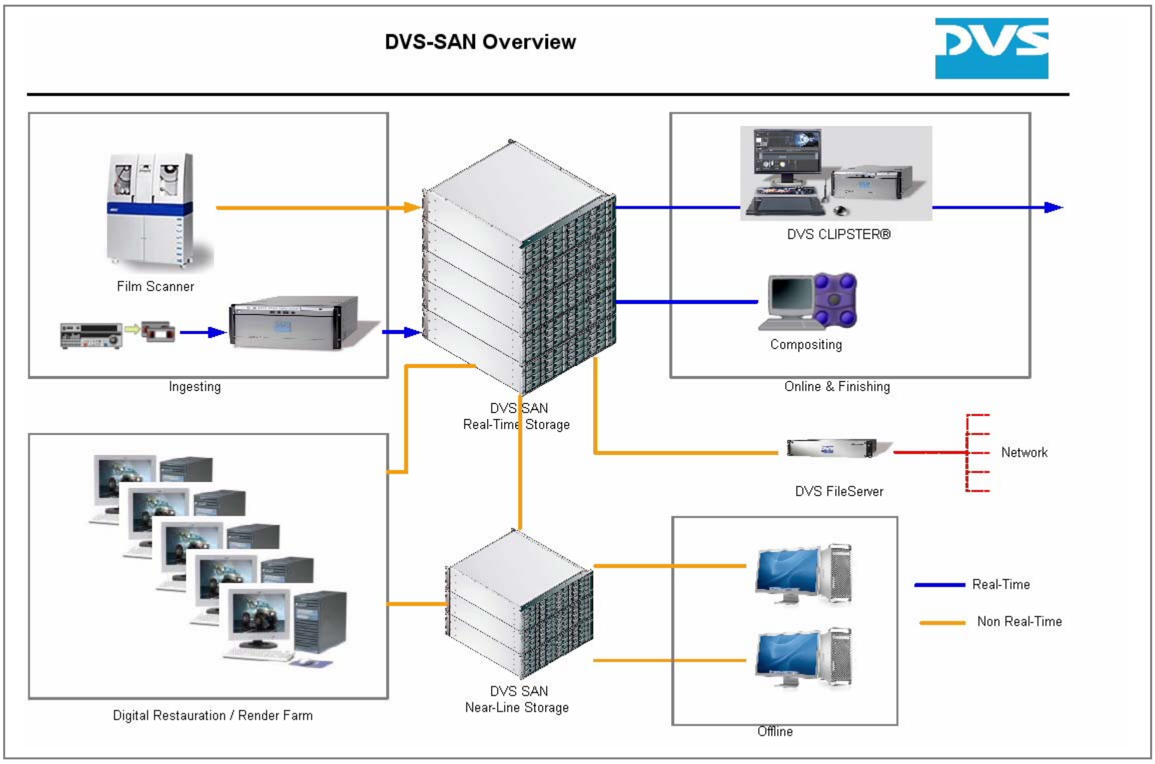

Image 13: Real-time and near-line SANs as center points (Copyright: DVS)

Key Facts: Application Examples
These application examples might give you an idea on how to set up a SAN in your facility – or how to improve the setup. The DVS-SAN is a flexible unit: define your requirements and your goals – and the DVS-SAN will undoubtedly fit in your requirements!

6. Innovative Data Management – a Solution by DVS

Not only data storage is essential in the day-to-day running of a postproduction facility, but effective data management as well. The following paragraphs deal with a completely new concept in SAN management. The application Spycer™ is an integral part of DVS-SAN. With this innovative data manager postproduction businesses worldwide can move to a higher level of efficiency.

6.1 Summary of Requirements

So, in review, there are several important facts to consider when implementing a SAN in your postproduction business. A quick re-wrap will prepare you for the next steps. First, **collaboration**.
A SAN allows for the collaboration of working groups or complete facilities. DVS-SAN is configured for access of many users using the data. Therefore a high and sustained data rate is an absolute must for a DVS-SAN.

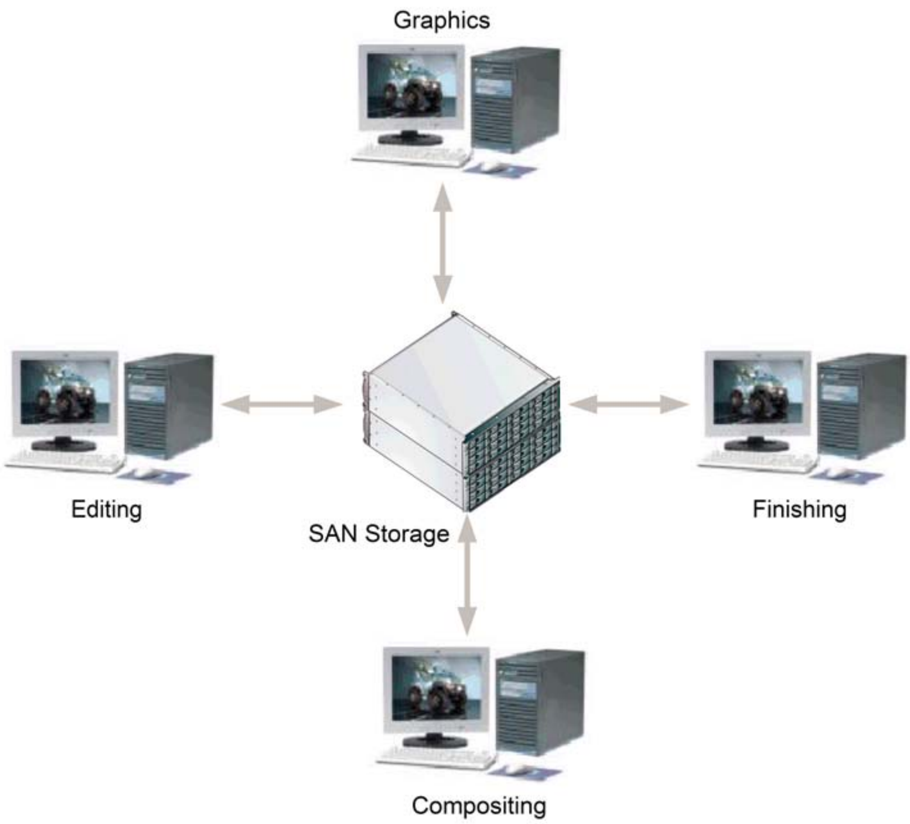

Image 14: Collaboration – everyone can work on the same material in the different stages of a production (Copyright: DVS)

Secondly, **customizing the system**. Customer customizations needs to be reflected in the SAN design, as the SAN is the information hub for several workflows, or for all workflows of a facility. The DVS team works closely together with the customer to collect all requirements that are then turned into a SAN configuration. Different hardware components from different OEM manufacturers are an issue, as well as insight knowledge of the file systems used. Topped with DVS own software components a DVS-SAN results in exactly the right configuration. Only an inside knowledge of video and IT engineering can result in the best possible performance and reliability.

Scalability is another main issue. Requirements for storage size and speed might change just a few weeks after the initial purchase of the unit. In the configuration stage future expansion must be taken into consideration. DVS-SAN is easily expandable in size, data rate and the number of concurrently accessing clients. Therefore the technology can grow as the business grows.

Its **flexibility and independence** is second-to-none. Usually a workflow includes different systems from different manufactures running on diverse operating systems, using different media file formats. The used SAN file system allows connected Irix®, Linux®, Mac and Windows® clients to the DVS-SAN and to use the storage concurrently. For each workstation the SAN will look like a storage volume of its own file system.

Image 15: Independence – No need to choose a specific operating system (Copyright by Apple®, Linux® and Microsoft®)

And last but not least, **reliability** is fundamental to any SAN system. This includes accessibility (low or even zero down-time), reliable data rates, as well as data protection. Utilizing fail-over configurations, a system with no single point of failure can be achieved. But also a fail-over for certain critical parts of the system e.g. the metadata server can be offered.
The data itself is RAID 5 protected. Even a disk failure will not result in a loss of data. After recovery the data will be recreated on the hot spare or exchanged disk and the SAN is still accessible with the same speed and data as before.

With the SAN in its final configuration, fitting the requirements perfectly, the next step is to think about how you can manage all the data that will be stored there. DVS can provide an innovative solution – Spycer™. The intelligent content management system manages hundreds of terabytes of data, possesses useful search functions and browsing tools and a sophisticated defragmenter (See also chapter 4.3.2 Real-time storage). The next paragraphs will give insight into fragmentation, file handling as well as content and network technology.

6.2 Spycer™ – the innovative Data Manager

Fragmentation of image sequences: In a digital intermediate process or digital cinema production, clips are stored as uncompressed file sequences. This means that a large amount of files are stored on a SAN. Most SAN solutions do not provide efficient defragmentation software for these file sequences. Common defragmentation applications just defragment single files and are "not aware" of the relationship between files. For real-time play-out, it is not enough to have just single defragmented files. The whole sequence of files must be stored contiguously in large blocks on the storage and not scattered all over it.

File Format	Defrag	Size	Resolution	Color Depth
DPX	99%	1.02 GB	1920 x 1080 Pixel	10
DPX	0%	2.32 GB	1920 x 1080 Pixel	10
DPX	3%	3.56 GB	1920 x 1080 Pixel	10
DPX	100%	395.90 MB	1920 x 1080 Pixel	10
YUV8	100%	197.75 MB	1920 x 1080 Pixel	8
DPX	100%	1.37 GB	1920 x 1080 Pixel	8
DPX	100%	594.14 MB	1920 x 1080 Pixel	8
DPX	100%	594.14 MB	1920 x 1080 Pixel	8
DPX	100%	534.73 MB	1920 x 1080 Pixel	8
DPX	100%	475.08 MB	1920 x 1080 Pixel	10
BMP	100%	183.94 MB	1920 x 1080 Pixel	8
BMP	100%	282.75 MB	2048 x 1556 Pixel	8
DPX	100%	1.16 GB	1920 x 1080 Pixel	10
DPX	100%	1.36 GB	1920 x 1080 Pixel	10
BMP	100%	995.97 MB	1920 x 1080 Pixel	8
DPX	100%	942.24 MB	1920 x 1080 Pixel	10

Image 16: A data manager must show the current defragmentation status of scattered sequences (Copyright: DVS)

Another important point is, that the storage data rate should not drop when a defragmentation job is launched and running as a background process.
A defragmentation tool has to be aware of the current workload.

Spycer™, the new content management software from DVS, is able to defragment any file and to analyze and de-scatter file sequences. A monitoring process updates the current defragmentation status of every sequence and keeps the user up-to-date. Spycer™ is "aware" if a workstation is under heavy load for real-time play-out and halts the defragmentation process if necessary.

File handling: Users require dependable copying and renaming processes for such masses of data. Fast copying mechanisms should be easy to use for the creative staff at their workstations. This means that drag'n'drop procedures must be available instead of cumbersome command line tools. DVSCopy, a fast copy tool that avoids fragmentation, is integrated in the GUI of Spycer™. Users see two browsing windows which show either source or destination storage of the files.

Image 17: Browsing windows with two directories in parallel provide for easy drag'n'drop use (Copyright: DVS)

With a built-in renaming function for file sequences, it is very easy to change the naming index or the whole naming pattern. Spycer™ goes through all files and renames them according to the given pattern automatically.

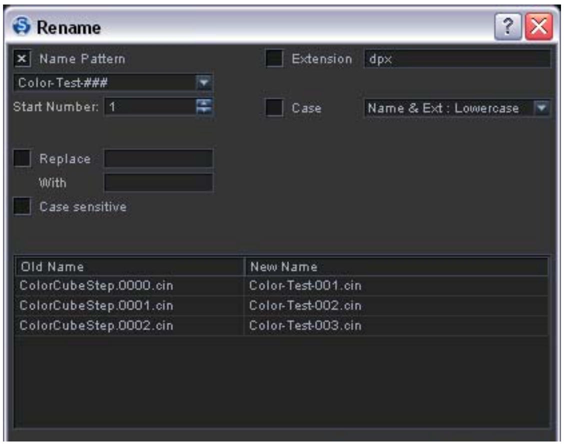

Image 18: File sequences require a good renaming function (Copyright: DVS)

Finding the content: project-based file hierarchy vs. search engine and project metadata.
For data handling and tracking, users need intelligent software which monitors every change made in a directory. It is not always possible to find an ideal hierarchical directory structure for a project in which everyone can find requested material on a big central SAN easily. In order to find a specific file, instead of browsing through several directories manually, it is faster to use a search tool which provides specific attributes for the search query. This is a new working paradigm. In Spycer™ a powerful retrieval engine is implemented. This holds important metadata attributes to localize data fast on the whole network.
Every Spycer™ application monitors its own volumes and provides automatically extracted metadata to the SpycerNet. A distributed content management network consisting of individually installed Spycer™ applications, is an ideal tool for the postproduction, where content is spread over different storage systems.

Image 19: Stored metadata help to find the right content. (Copyright: DVS)

Network technology: Client-Server System vs. Distributed Content Management-Network
It is hardly ever the case that a central SAN or NAS is the only storage where all data is stored and can be accessed from everyone involved in the project. In the majority of cases, postproduction facilities deal with several islands of storage solutions for security, performance or economical reasons.
In a traditional IT structure there exists a central storage and a central database where all information is administrated. The problem is that a lot of media files never get recorded in the database. This is especially true if an ordinary client-server structure forms the basis for a content management system (CMS). Often a client software provides just a GUI for the database of the CMS. The user has to enter all data manually. For automatic metadata extraction, files have to be stored on the central system where background processes running around the clock can continuously update the database. Such systems are often limited in scalability and performance due to a single server and database being responsible for all files.

What if several databases and active software applications are used on every workstation? Imagine a distributed content management solution where every workstation automatically provides its information to a shared network. Every file gets a record in one of the databases which is available to all others in the company network. This is how a scalable system should work. According to the rights provided to an individual user, the user can find any media file in the network effortlessly by using common metadata information.
Spycer™ works exactly that way. Via SpycerNet several Spycer™ applications communicate with each other for browsing, data retrieval and file handling purposes. Metadata is automatically extracted and can be used for search and retrieval in the network. The distributed content management system grows with every additional Spycer™ application added to the network.

Rights management: in such a distributed content management network rights management is important. Scanning your own files and using available metadata from others on the network for a fast search function does not necessarily mean that other users have the right to view your files! For this, Spycer™ provides rights management for all metadata and files.

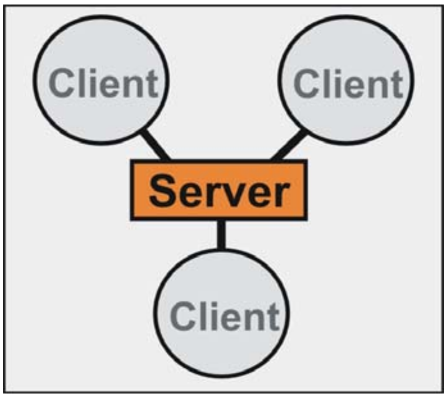

Image 20: Client-server system: a centralized approach (Copyright: DVS)

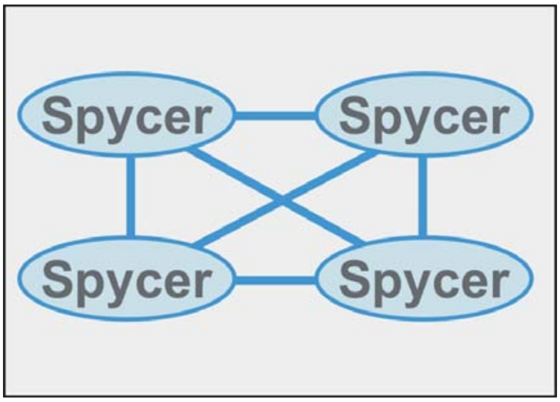

Image 21: SpycerNet – a truly scalable distributed content management (Copyright: DVS)

Key Facts: DVS-SAN and Spycer™
The intelligent configuration of DVS-SAN deals with any kind of postproduction workflow requirement. Collaboration on a customized system and the scalability of DVS-SAN offer the best possible performance and reliability. Its flexibility and independence are enhanced with a new innovative application: Spycer™ is able to analyze and defragment large clips. A monitoring process updates the current defragmentation status of every sequence and keeps the users up-to-date.

7. DVS: Innovations for the Film and Broadcast Industry

For more than 20 years DVS with its headquarters in Hanover, Germany, has been a leading manufacturer of high-performance digital video products for film, TV, postproduction and R&D. DVS offers a wide range of premium software and high-end video turnkey workstations – from disk recorders and storage systems to conforming and finishing systems for Digital Intermediates. DVS software and hardware are recognized as a reference for quality, reliability and performance.

In addition to its award-winning DI workstation CLIPSTER® DVS has presented another milestone in its software development: at NAB 2006 DVS unveiled the content management system Spycer™. In 2006 a new line of disk recorders came to market: Pronto2K.2 and ProntoHD.2 enable real-time multi-resolution capture and play-out combined with full conforming. The company has a large OEM customer base, too: numerous leaders in the industry integrate DVS' powerful video I/O boards into their own products. Moreover, DVS has teamed up with ARRI, Kodak, Imagica and The Mill.

A complex technical setup like a SAN needs a competent partner with video and storage know-how. This starts with the configuration of the system and is continued with the installation on site. A plan for setup and replacement must be in place in case of a component failure. For the best possible support, DVS offers advanced replacements of the parts, spare parts stored at the customer site and has a worldwide service partnership for on-site replacements. Different service packages allow up to 24/7 phone and remote diagnostics support, covering every time zone and on-site maintenance visits.

With more than 20 years of experience and high-quality products DVS hardware and software solutions are strongly sought after amongst industry leaders.

Key Facts: DVS Quality
The DVS quality and experience is well-known and appreciated in the industry.
Flexibility and innovative concepts from DVS have provided the industry with major milestones.
On top of that, DVS offers support packages to suit all user requirements.

8. Conclusions

From celluloid to digital data, the history of film and video postproduction has undergone rapid changes. Today's requirement for high data rates and data quantities impose interesting challenges for the industry.
A SAN is an excellent solution to handle the large quantities of data and the requirement for collaborative workflows in the postproduction facilities. A SAN is a Storage Area Network. SANs are easily scalable and expandable in regard to speed, size and number of connected clients.

Depending on the individual requirements in a facility, different configurations have to be taken into account. DVS-SAN offers a wide range of possibilities. This can start with a simple near-line storage attached to two clients, up to configurations with dozens of clients connected to huge amounts of storage with several real-time 2K data streams in parallel. Both its hardware design and software setup enable the facility to handle the amount of data simply and reliably.

There are quite a few questions to be answered when considering investment in a SAN. Of course, it is essential to know how much storage is needed. But there is more to it than the famous money versus storage calculation. And there are more questions to be asked and answered: How many clients will access to the SAN? Is the SAN in the need of a real-time storage or a near-line storage? What is the best way to organize the data? Are there different operating systems in use? Is there any protection against failure – and what support is available when needed?

DVS-SAN has unrivaled flexibility. This is enhanced with a new innovative application: Spycer™ is able to analyze and defragment large clips. A monitoring process updates the current defragmentation status of every sequence and keeps the user up-to-date.
DVS quality and experience is renowned in the industry. Flexibility and innovative concepts from DVS have provided major milestones to the industry. On top of that, DVS offers various support packages to suit all user requirements.

Undoubtedly, a SAN is a perfect solution for many postproduction facilities. Sustaining highest data rates and real-time capabilities is what is required. Faster network technologies will allow a wider use of distributed storage and real-time NAS storages in the future.
Nevertheless, DVS-SAN is a future-proof investment since it seamlessly integrates with NAS, DAS (Direct Attached Storage) and data networks.

9. References

Apple: http://www.apple.com/pr/products/ and
http://www.apple.com/downloads/dashboard/games/applelogo.html

DVS: http://www.dvs.de

Hitachi: http://www.hitachi.com/

Linux®: http://www.isc.tamu.edu/~lewing/linux/

Raid: http://searchstorage.techtarget.com/gDefinition/0,294236,sid5_gci214332,00.html

RoHS: http://europa.eu.int/eur-lex/pri/en/oj/dat/2003/l_037/l_03720030213en00190023.pdf

Seagate: http://www.seagate.com/

Windows®: http://www.microsoft.com/presspass/gallery.mspx

Wikimedia: http://commons.wikimedia.org/wiki/Image:Film_reel_and_film.jpg

10. Glossary

2K
Abbreviation for resolutions of image and video, e.g. 2048 x 1536 and 2048 x 1556 pixels. This resolution is recognized as the minimum resolution that is suitable for cinema presentations. In order to obtain other aspect ratios or to be compatible with technologies like HD-SDI, further formats like 2048 x 1080 and 2048 x 858 are common. As 2K is used for film production mainly, the color coding is mostly RGB 4:4:4, which gives a better quality than YUV 4:2:2.

4:2:2/:4
This is a term commonly used for a component digital video format. It represents a certain ratio of sampling frequencies used to digitize color difference components (Y, R-Y, B-Y) of a video signal and the luminance. 4:2:2 describes that for every four samples of Y, there are two samples each of B-Y and R-Y, increasing chrominance bandwidth in relation to luminance compared to 4:1:1 sampling.
4:2:2:4 is identical to 4:2:2 with a further channel for key channel that is sampled four times for every four samples of the luminance channel.

4:4:4/:4
Quite similar to 4:2:2 with one exception: luminance and chrominance are sampled with the same rate. 4:4:4 is often used with RGB color coding, but YUV color coding in conjunction with 4:4:4 sampling is also common. 4:4:4:4 is similar to 4:2:2:4. Here as well, for every four luminance samples, the color and key channels are also sampled four times.

4K
Abbreviation for an image and video resolution of 4096 x 3112. This resolution is recognized to be of the same quality like a high-quality cinema presentation. To obtain other aspect ratios or to be compatible with technologies like HD-SDI or DVI, further formats like 4096 x 2160 and 4096 x 1714 are common. The 4K resolution is higher than the one of a HDTV signal. As 4K is used for film production mainly the color coding is mostly RGB 4:4:4 which gives a better quality than YUV 4:2:2.

ADIC
A company providing data storage solutions, Advanced Digital Information Corporation.

ATA
Advanced Technology Attachment. See IDE.

Cache
A cache is an especially fast memory. It is a collection of duplicated data values stored in a memory.

Centaurus®
Centaurus® is the industry standard for high-end uncompressed video I/O hardware. Based on PCI-X bus architecture this DVS video I/O board combines the technology of DVS' SDStationOEM and HDStationOEM. Centaurus® is equipped with RS.422 remote control, wordclock, GPI interface, and a real-time hardware mixer. The board is ideal for compositing, title generation, virtual studio, color correction, and simple video I/O. Its successor, Centaurus II, will be on the market in late 2006 and is then available in a PCI-X 133 and PCI Express version.

Cineon®
This is a file format that was specifically designed to represent scanned film images.

Client
A computer system that wants to access a service – sometimes a remote one – on another computer is called a client. Typically this happens within a network.

Client-Server Architecture
Network structure which separates server applications from client applications. A central server manages all data for different clients and provides them with the required data. The system's scalability depends on the server performance and the expandability of its hardware resources.

CLIPSTER®
CLIPSTER® is a turnkey solution by DVS. It is a one-stop Digital Intermediate solution for conforming and finishing uncompressed SD/HD/2K/4K data in any workflow. Moreover, this DI workstation carries out real-time effects, enables multi-resolution and is an open platform. CLIPSTER® offers stunning hardware power and innovative software for unrivaled flexibility and can be used in any video or film post-production environment. The high performance is to be seen in its real-time effects with up to 2 x 2K RGB 12 bit, its real-time playback of 4K RGB 10 bit DPX file sequences and its support of multiple video formats with real-time converting. Additionally, CLIPSTER® can handle real uncompressed video up to 4K RGB 16bit and runs real-time effects in 16 bit, with original native content being used for real-time processing. CLIPSTER® is an open platform: the Windows® XP workstation captures directly to NTFS and it possesses real-time support of graphic file sequences like DPX, TIFF, Cineon®, TGA, BMP, etc. Of course, an OpenFX plug-in interface is part of CLIPSTER® as well.

CMS
A Content Management System is a software that helps storing content and tracking changes made by users. It supports the organization of content via a database.

Color space
This term describes the color range between specified references. Normally, references in television are quoted in the following way: RGB, Y, R-Y, B-Y, YIQ, YUV and Hue Saturation and Luminance (HSL), and XYZ. Common for these color spaces is that three values are used to represent a color. In print, Cyan, Magenta, Yellow and Black (CMYK) are used. It is possible to convert images between these color spaces, but due to the accuracy of processing involved care is required. When operating across the media – print, film, TV, as well as between computers and TV equipment – conversions in color space are needed.

Component
Each color of an RGB or YCbCr signal is transmitted via an indivual cable. This would be called a component color signal in opposed to a composite signal, where the three components are transmitted via one cable.

Compositing
Simultaneous multi-layering and design for moving images. Modern designs often use different techniques in combination, such as retouching, rotoscoping, painting, keying/matting, digital effects and color correction as well as multi-layering in order to create complex animations and opticals in promotions, title sequences, commercials as well as in program content. Besides the creative element there are other important applications for compositing equipment such as image repair, glass painting and wire removal - especially in motion pictures. The quality of the finished work, and therefore the equipment, can be crucial. This is especially true where seamless results are demanded. For example, adding a foreground convincingly over a background - placing an actor into a scene - without any telltale blue edges or other signs that the scene is composed.

DAS
Direct Attached Storage. A storage unit directly attached to the device recording the data.

DDR / digital disk recorder
Systems that record video or audio programs on one or more hard drives. They are mostly used in broadcast or radio broadcasting when editing or recording is required. The benefit of these systems: they offer immediate access to the material that was recorded before, without requiring pre-roll/post-roll or expensive maintenance of tape heads.

DPX
Abbreviation for Digital Picture Exchange. This file format can be found in digital film work and is considered an ANSI/SMPTE 268M standard. DPX files can store image data and additional metadata in their file header.

DVS-SAN
DVS-SAN is a high-performance central storage system for uncompressed video and digital film. It has been designed for use in the film and HD postproduction environment, where large amounts of data must be accessed in real time by multiple workstations. DVS-SAN meets the special demands of digital intermediate work and HD projects requiring ultrahigh data rates and fast access times. With DVS-SAN, several workstations can access the same data, concurrently and in real time, eliminating the need for copying and exporting. The system is upward-scalable to hundreds of terabytes and so offers enough storage capacity for several film projects.

EDL
Edit Decision List, a file that describes how video or film sequences shall be assembled. In general it contains four columns with time code information. Two columns define the source position and two the destination position of the clip. EDLs can contain additional information about special effects and others.

Ethernet
Ethernet is a network technology for data transmission. A star-topology with twisted pair wiring is the most popular form. Common data rates are 10 Mbit/s (Ethernet, 10 Base-T), 100 Mbit/s (Fast Ethernet,100 Base-T), 1000 Mbit/s (Gigabit Ethernet, 1000 Base-T) and 10,000 Mbit/s (10 Gigabit Ethernet).

Fail-over
This term describes the capacity to automatically switch over to a redundant system, network or computer server, when a failure or abnormal termination of the active system, network or server occurs.

FC-drives
See Fibre Channel. These drives use the copper version of the Fibre Channel interface with a SCSI protocol. The maximum data rate is 4 Gbps.

Fibre Channel
A Fibre Channel is a universal, high-speed data link that can handle up to 4 Gb/s on a fibre optic cable. It originates from computer technology but is used in the video industry as well. The industry shows great interest in products using Fibre Channel, e.g. so that hard disks can be connected. Fibre Channel can be transmitted optically via optical fibre or electronically via copper cable.

File System
When storing and organizing computer files and their accompanying metadata, a popular method to use a file system. A file system might possibly have a storage device (e.g. hard disk) and then maintaining the physical location of the files is of importance. The file system will translate the file name used by the user to the physical address on the storage device. Another option is that the file system grants access to data on a file server – then acting as clients for a network protocol. File systems might be virtual, too, and then only exist as an access method for virtual data.

Format
(1) The size, resolution, aspect ratio, color space, bit depth, format rate, etc. for a given image.
(2) The file format for a given image.
(3) The physical medium (such as film, video, etc.) used to capture or display an image sequence.
(4) A multitude of additional variations and subcategories of the first three definitions.

Fragmentation
(Data) fragmentation occurs when a piece of data in memory is divided into several parts being physically far apart. Generally, this is the result of attempting to insert a large block of data into several small free spaces on the storage.

GUI
Graphical User Interface. An interactive graphic displayed on a screen, being a means of operating a software.

HD
High-definition. Sometimes used as a short form of HDTV.

