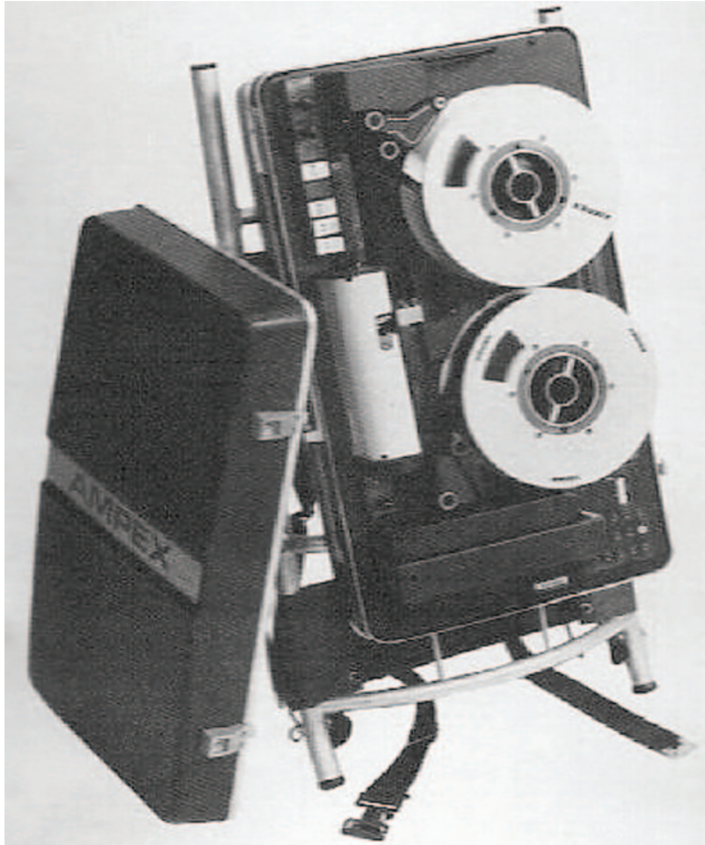


The first portable VTR. This two inch VTR was introduced by Ampex in 1967. It was strapped to the back of an operator. This task was best for 'newbies' as they needed to be young and have a strong back. This machine weighed 35 pounds, could accept a 20 minute reel of two inch tape, and its battery would need to be recharged after that 20 minute tape was used up. Like many portable VTRs that came after it, this machine only recorded. The system cost $65,000 and was the first system aimed at one-person field production. Ampex also offered a camera that mated with it. The picture on the right is of a still surviving and operating 3000.

VTRs

It's strange to see the demise of videotape as it was once a real growth format. We had one tape format at first, in 1956, with a few more developing in the 1960s, 1970s and 1980s; at the turn of the century there were more than a dozen different tape formats. A large part of the reason for all those formats was that most of the later formats were digital. That means that audio and video are recorded onto tape as data. Digital data requires lots of bandwidth, and compression is used to limit that bandwidth to make reasonably priced VTRs. ($20,000 to $60,000 units were considered "reasonable" for broadcast). Different manufacturers had differing philosophies on video and audio compression and the best tape size, hence all the formats.

The VTR truly became portable with the advent of the ¾ inch tape VTR, known generally by the Sony brand name Umatic. The machine shown here (on the left) was known as the VO-3800 (the box on the right is its optional power supply). This VTR was originally intended for industrial video use, but was adopted by innovative broadcasters for ENG use. It weighed less than 25 pounds, but with ENG cameras of the day, and battery belt, the photographer was lugging over 50 pounds of gear. Many a photographer that was used to shooting 16 mm film soon developed back problems.

Cameras, and especially VTRs, have seen a reversal in technology flow. It used to be that RCA, Sony, and Panasonic would develop technology for their high-end product offerings, such as the ones for broadcast, and that technology would eventually trickle down into consumer gear. Now the opposite is the norm.

Many of the professional VTRs today are based on tape sizes developed for the consumer. The most well-known is the consumer Beta tape format that lost out to VHS in the 1980s. The Beta tape morphed into the high-end Sony Betacam VTR family. VHS did the same thing, becoming Panasonic's MII, a format no longer in use. The tape and some of the tape transport technology were the same, but the electronics were definitely high-end, as compared to the consumer gear.

The machines here were once high end mainline VTRs used for recording and replay. These machines used one inch wide tape. They produced higher quality video than two-inch VTRs. They also played at speeds up to two times real time, down to stop motion, even slow reverse. One inch machines were the workhorses through the 1980s and into the 1990s.

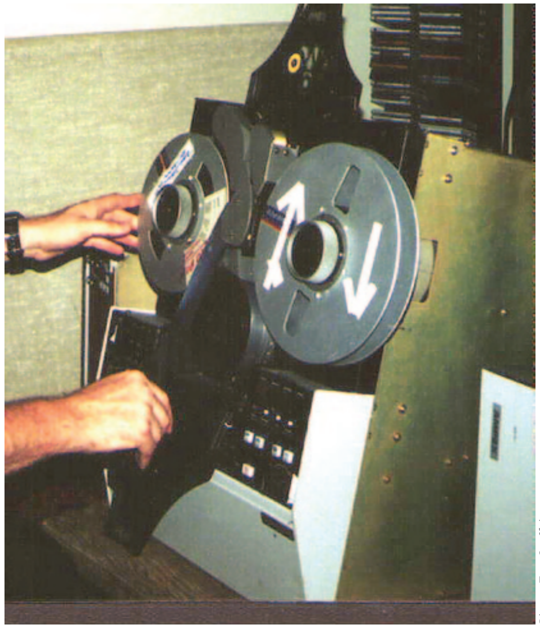
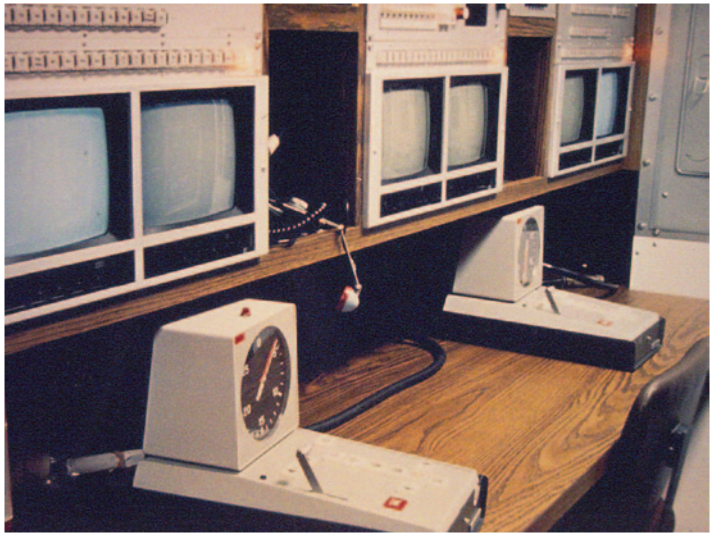


Photo: Tim Stoffel

The most common machine used for replay, especially through the 1980s was Ampex's VPR-2. The base machine is shown here. Notice the taped arrows on the closest tape reel, which is the take up reel. Many replay operators would manually spin the take up reel to get back to the replay start point and then continue to turn the reel for the actual replay. Many operators thought they had better control and could affect more finesse over playback this way. In fact, many would show up for work with their own specially modified take up reel, complete with a knob drilled into the reel.



The basic look and function of the slo-mo controller hasn't radically changed since the introduction of the Ampex disc based slo-mo system in 1967. There is a photo of that system in Chapter 1's Pre-HD section. The controllers seen here were on a truck in the 1980s.

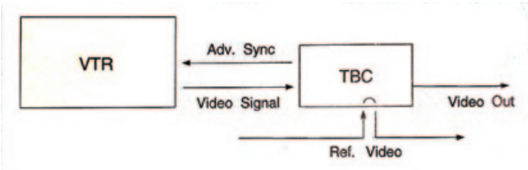
Here is a CBS owned truck in the early 1980s. This trailer only housed VTRs and an Ampex disk-based slo-mo system. As can be seen here, the transition from 2 inch VTRs (machine in the foreground) to one-inch VTRs is occurring.

Portable one-inch VTR. This one was manufactured by Nagra, a well known audio recording company, which produced this machine in conjunction with Ampex. As the technical OEM, Ampex branded and sold it. Ampex offered an earlier version that was problematic, the VPR-20.

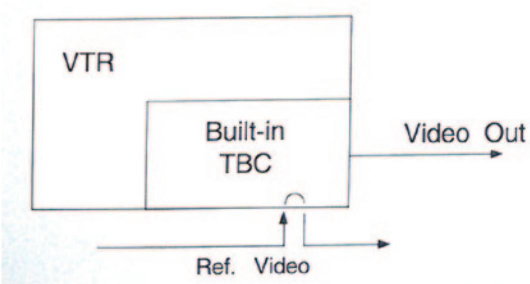
This was the Ampex original portable one-inch VTR, the VPR-20. It was introduced in 1979. It was aimed at the EFP market."

Sony's portable one-inch machine. Notice that the two tape reels are atop one another. These coaxial tape reels had the take up reel on top. The Ampex VPR-20 had the same configuration.

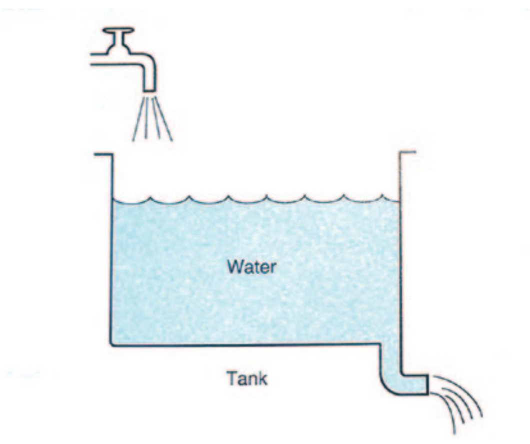
Most people do not realize that an early beach head for digital video technology was associated with the VTR. VTRs recorded and recovered video information using a mechanically spinning head. As such, these devices were continuously running a little faster or slower than desired. This velocity error, as it was called, would play havoc with video timing, especially for the subcarrier embedded in the video, carrying the color information. This was caused by the fact the scanner and capstan can never exactly duplicate the exact conditions found when recording, or match the playback video to "house" video, which is used as the external video reference. To even out the playback, the video was converted to digital, marched into memory at the varying rate it was arriving in from the VTR, and marched back out at a steady rate that matched the required timing of the facility housing the VTR. This was in the early 1970s; the device was known as a Time Base Corrector or TBC. Before this, elaborate banks of electronics with lots of variable capacitance were used. These pre-TBC subsystems were large and very finicky.



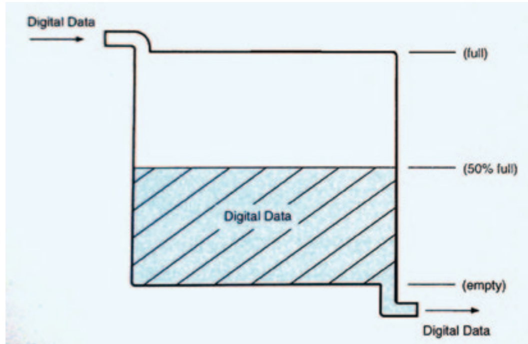
Originally, TBCs were separate devices that received varying video from VTRs and stabilized it.



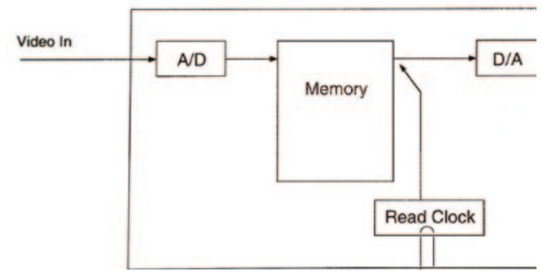
Eventually, the TBC was incorporated into the VTR and became an integral part of the unit.



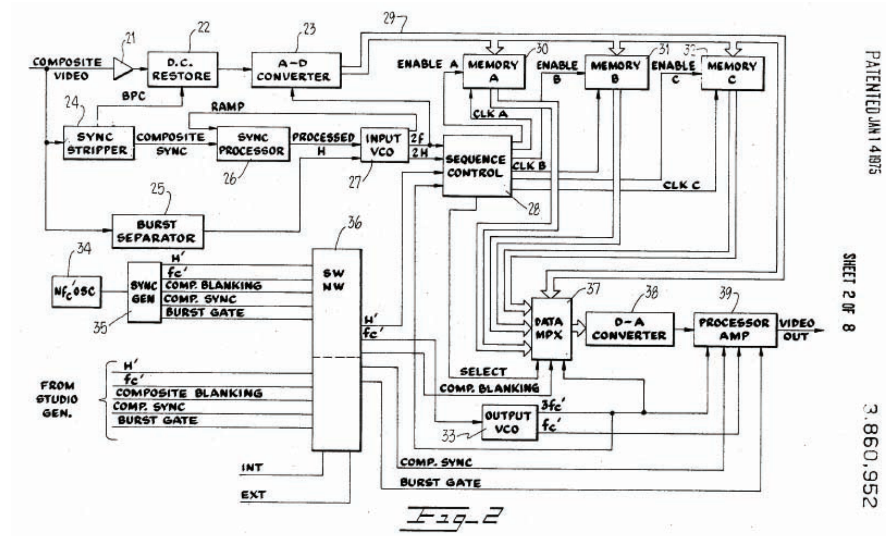
A TBC can be thought of as a water holding tank where the faucet into the tank is constantly being opened and closed to varying degrees. But the water out of the bottom of the tank will be constant, unless the tank is allowed to completely empty.



The same analogy can be extended to a digital system. The water represents digital data, the size of the tank represents the amount of memory. Data continues to flow out of memory at a steady rate, unless the tank is allowed to empty because the data inflow is too low, known as a buffer underflow. If the tank, or memory, is allowed to overflow, that data is lost.



From our tank analogy to an electronic depiction: video in at varying rates, video out at a steady rate determined by the video reference signal. This basic process is still what happens today when video is played out of any video server. Data is read off the spinning disks in the server, and those disks have angular velocities (the rate at which they are spinning) that are not constant. So the TBC process must still be performed.



Consolidated Video Systems sprang up in Sunnyvale, Ca. just as Silicon Valley was gaining its footing. It applied for a patent on the Time Base Corrector (TBC) in 1973. It offered its first TBC in 1975. Here is a block diagram that supported claims in their patent application. As can be seen, near the beginning and end of the TBC food chain are converters to get the video into and out of the digital domain. There were two great bottlenecks in technology that governed early TBC development; the speed that electronics of the day could operate at, and the cost and size of the memory. Early on, the TBC was a stand alone box, at first many rack units tall and eventually only a single rack unit high, with capabilities several orders of magnitude greater than the first ones. Within 15 years, TBCs were starting to be incorporated into VTRs. Frame Synchronizers and many effects devices today still must perform the same steps seen here, only with ASICs, embedded processors, and software—or a combination of the above.

The memory used for these digital TBCs was very expensive and took up a lot more space back then, so generally only two or three horizontal lines could be stored. Eventually, memory got smaller and cheaper, and the number of horizontal lines that could be stored grew, until a whole frame, or even multiple frames of video could be stored. When that happened, the frame sync was born, and shortly after that, the digital video effects device was introduced. The rest is history, as they say.

VTRs were always high maintenance devices. From the days when they weighed 1400 pounds each, and occupied the space that two side-by-side refrigerators would command, to the latest devices that still actually play tape, they by nature, have required continual and periodic maintenance. There's a thermodynamic term called entropy. Besides predicting the long term slow loss of heat throughout the universe, it also claims that over time all systems dissolve into disorder from order. On a smaller scale, this accurately describes the VTR.

The very nature of recording and playing back tape slowly erodes all in the tape path. From guides and rollers to capstans, audio heads, and especially video heads with their mounts on rotating scanners, friction is always at work. In fact, tape and friction have caused an almost constant gnashing of teeth right up to this day.

Tape is a strip of Mylar, coated with a secret blend of iron oxide and other materials, some of which lubricate surface contacts and mitigate the friction between tape and its path. Sometimes, manufacturers got the amount of friction wrong in some batches of tape. You see, some friction is required, or video heads would tend to coat up with oxide and quit recording or playing back information on the tape, otherwise known as "head clog". Too much friction and the tape path would be ground down too fast.

Another confounding fact was that the video heads on the rotating scanner wouldn't wear down height-wise quickly. They would deform in a way that reflected the aerodynamics of their path in contact with the tape surface. This would set up turbulence affecting the contact of other heads on the scanner following. Users would think the heads were worn down, when they simply needed to be reshaped. A few seconds of running a "cleaning tape", which had a highly abrasive surface, would fix that. There was controversy about the tape's grinding everything else in the path, but being able to recover the use of the rotating scanner, usually replaced as a single unit at a cost of over a thousand dollars, usually trumped the other considerations.

This upper rotational scanner could have a dozen or more heads because of replay and editing requirements. It wasn't uncommon for some trauma to break a head; not only would that particular head not do its job, it could very well start shredding oxide off the tape. These heads were often under tremendous stress also: the record heads required many amperes of current flowing through them to re-align the magnetic particles in the tape oxide. In fact, the current was high enough that if tape wasn't in constant contact with the record heads, they would burn themselves out in short order.

Even with these limitations, you still find VTRs and their associated tape on trucks. While over 90% of trucks today have video servers of some flavor, many still have a sampling of VTRs. Sony's digi beta format is still dominant, and you will find well over 50% of trucks still carrying one or more of these. On a third of today's trucks you'll still find the venerable analog beta (both sp and sx formats), HDCam, and DVCPro tape formats. other production folks will want to leave with material in one of those formats.

10 years ago, this tape or replay compartment, which consumed an entire trailer, was used for Monday Night Football. Here it is being used for the Indianapolis 500. Its usual complement was 16 VTRs and a few channels of EVS server storage. Today it is not unusual to have over three times the replay capability of this truck.

A common format still in use is the Digital Betacam. It uses a mild 2:1 JPEG compression scheme. It records 4:2:2 component video at 10 bits. The data rate to tape is 127.8 Mbps. Digital Betacam uses ½ inch tape, and moves it at 96.7 mm/s. Optionally, this machine can play back analog Betacam tapes even though they are completely different formats on tape, with the power of advanced processing and a couple of mechanical tricks regarding the scanner.

The DVCPro is based on the consumer DV format, but it is beefed up and intended for ENG applications. Many mid-size players used it as a production format due to its relatively low cost. This format comes in a number of options, generally based on the rate that data is recorded onto tape. All 3 types use 6.35 mm thin metal particle tape. Regular DVCPro uses a 4:1:1 sampling structure at 8 bits. It pulls tape at 33.8 mm/s and records data on tape at 25 Mbps. It uses 5:1 JPEG intra-frame compression. This is machine can play back tape at faster than real time (4 times). It records two 16 bit audio channels sampled at 48 KHz. DVCPro will playback DV-CAM tapes.

DVCPro 50 lays down data at 50 Mbps but at double the tape speed of DVCPro (67.6 mm/s). It has a 4:2:2 video sampling structure at 8 bits. It also uses JPEG intra-frame compression, but has only a 3.3:1 compression ratio. It records four 20 bit audio channels also at 48 KHz.

DVCPRO HD, also known as DVCPRO100, is an HD format that also uses 4:2:2 color subsampling and is backward compatible with earlier DVCPRO formats. The video data rate ranges from 40 mbps for 24 frame video up to 100 mbps for 60 frames. DVCPro uses a smaller raster size than HD: 960x720 pixels for 720p, 1280x1080 for 1080/59.94i and 1440x1080 for 1080/50i. This is also known as "thin raster". On playback, a DVCPRO player converts video back to normal HD raster sizes. DVCPRO HD uses the same 18 µm track pitch as other DVCPRO formats. There is a longer playing variant, known as DVCPRO HD-LP, which doubles the recording density by using a 9 µm track pitch. DVCPRO HD wraps video and audio media into the MXF Op-Atom format. Camcorders often no longer have DVCPRO HD tape decks, but record onto P2 memory cards.

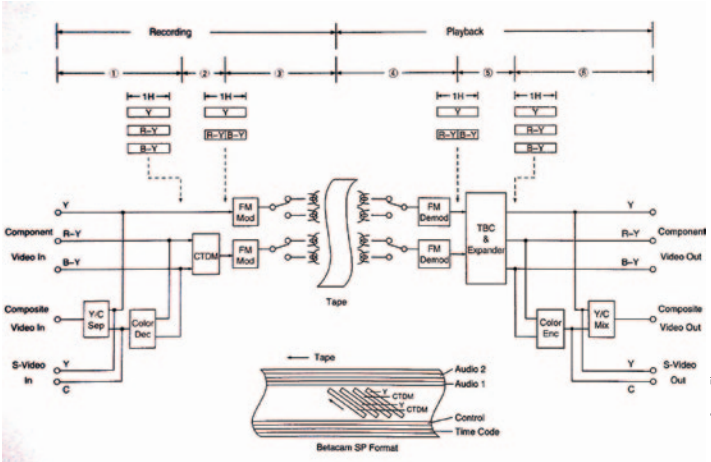


Image: Sony Electronics

Block diagram of a Sony Betacam VTR, a venerable workhorse for over 20 years. In many ways, servers today have the same basic architecture: an input side, storage medium in the middle, and an output side. FM was used to lay information on tape as it required much less bandwidth than other methods. Being a component machine, the Y, or black and white signal, is separate from the R-Y and B-Y color signals. Notice the internal TBC. The tape format at the bottom gives hints as to why so many tape formats continue on today as file formats. Even though as an analog format, and thus never carried into the file format domain, you can see that the heads laid video information as diagonal tracks across the tape. Alternating tracks carried Y, and then a CTDM track. CTDM stood for Color Time Domain Modulation. This is a technical way to say that both R-Y and B-Y were compressed in time and laid down on the same track. These tracks are interleaved with one another. We use that term today with file formats. While the other track names are obvious as to their purpose, the Control track carried a continuous series of pulses that control the tape speed.

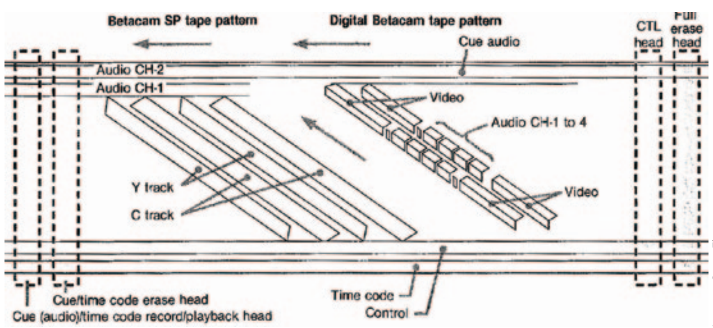


Image: Sony Electronics

As the transition from analog to digital began, there were machines that could handle both. The Sony DVW-A500 was one of those machines. While it couldn't record analog signals, it could play them back. This diagram shows the analog track layout (Betacam SP) on the left and the digital track layout (Digital Betacam) on the right. Here you can see that for the digital track format video, both black and white and color are interleaved with four channels of audio. The audio CH-1 track is eliminated as it is no longer needed. Also notice that the audio segments are laid down right in the middle of the tape. This is to minimize the possibility of losing audio data to tape damage, usually found along the edges.. While counterintuitive, missing data in audio is harder to correct than in video because of human perception. You should now be able to see how tape formats could be carried on into file formats that will never see tape. One final interesting thing about this machine is that the analog tracks were geometrically different from the digital tracks. The machine was able to recover either set of track patterns due to the incredible agility of the playback heads to be radically and dynamically bent, via piezoelectric means, to stay over the analog tracks.

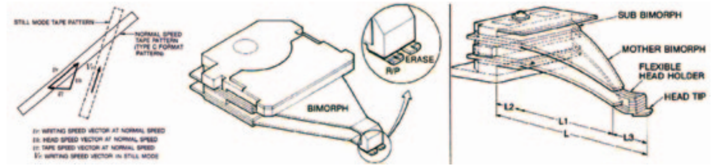


Image: Sony Electronics

Diagram showing piezoelectric head mentioned above. Sony called these Dynamic Tracking heads. This technology is what allowed VTRs to playback video at varying speeds, or even stills, and not display bands of noise in the video.

Migration Away From Tape


Photo: Corplex

10 years ago, this was the replay compartment of a high end HD truck. The VTRs shown here, at their heart, were SD machines, but modified to be able to record and playback HD. The input and output of these machines had external MPEG encoders/decoders. The data laid down on tape was simply MPEG compressed audio/video. Today, most HD VTRs do the same, except the compression and decompression are built into the machine.

The Betacam SX was the first "hybrid" VTR, as it included a disk drive for video/audio in addition to the standard VTR tape transport. It therefore, had both linear VTR and non-linear editor attributes. The base system has 2 hard drives (RAID-3 protected) that store 90 minutes of material. Over 6 hours of external hard drive storage can be added. Material on tape can be copied onto the hard disk at up to 4 times' real time. It uses MPEG-2 4:2:2@ML (I and B frames only—GOP of 2—more on this in the File Format section) 10:1 compression. The compression ratio is 10:1. The video is sampled at 4:2:2 at 8 bits. It also records four 16 bit audio channels sampled at 16 bits. SX passes the majority of the vertical interval through uncompressed. This format is intended by Sony for ENG applications.

Data is laid down on tape at 18 mbps. This low bit rate is intended, so SX data can move between machines through DS-3 type communications channels. Tape speed is 59.6 mm/s. SX uses ½ inch tape. The data rate and tape speed allows for a wide track pitch 32 microns. This wide track pitch also allows SX to eliminate the dynamic tracking head. Two staggered rotary heads cover an area wider than the track. The head receiving the strongest RF signal is used. Another neat feature of this machine is that it is possible to record and play back separate material simultaneously. Faster than real time play out can be achieved via tape play out but not via the hard disk.

HDCAM is another format from Sony, introduced in 1997, and it offered the first portable HD recorder. It uses ½ inch metal particle tape. Tape speed is 96.7 mm/s. Up to 124 minutes of HD material can be recorded. Like the SX format before it, much of the mechanics and electronics are based on Digital Betacam. Bit-rate to tape is 140 Mb/s. It samples video at 3:1:1 at 8 bits. It has a compression ratio just under 5:1 and lays data down on tape at 140 mbps. It records 4 channels of 20 bit audio sampled at 48 KHz.

In 2003 Sony introduced HDCAM SR, which recorded 4:2:2 or 4:4:4 at 10 bits and used a higher particle density tape. It has a video bit rate onto tape of 440 mbps. This format can record up to 12 audio channels.

Sony also had a DV entrant: DV-Cam. It sampled 4:1:1 at 8 bits and laid data onto tape at 25 mbps. Compression was JPEG at 5:1. It pulled tape at 28.2 mm/s and used ½ inch metal evaporated tape. Sony intended this format for the industrial user and DV-Cam can't playback DVCPro, because its track pitch is not as wide as DVCPro, but it can playback Consumer DV. The track pitch for DV Consumer = 10 microns, DV-CAM = 15 microns, DVCPro = 18 microns.

Now that tape is being replaced with other recording formats, many of the tape formats live on being recorded to optical discs, flash memory, and hard disks to be treated as files. Devices will stream DV-Cam over Firewire where it can be compressed and stored as AVI, Quicktime, or MXF files (again more on this in File Formats). DV-Cam media can also be stored uncompressed. These files usually have .dv or .diff extensions. DV-AVI is Microsoft's implementation of DV-Cam, as Quicktime-DV is Apple's version. Sony and Panasonic wrap DV-Cam into MXF containers. DV-Cam files stored on a computer are almost always stored in a DV, AVI, Quicktime, or MXF container file.

Sony, which no longer has a DV Cam tape deck, has an optical disk recorder called XDCAM which will record the DV-Cam format. At the Beijing Olympics, NBC had 180 XDCAM decks. XDCAM uses optical and solid state memory to store media. This format was also introduced in 2003. XDCAM camcorders, like other non-tape devices, are intended to replace traditional camcorder tape decks in news and production workflows. The decks can be brought back and used as hard-drives to ingest media into a post production workflow, or products from a number of vendors, will accept the various removable storage media, and ingest them in common storage for post production.

XDCAM uses 25 mbps DV encoding for SD, while MPEG2 IMX is used for SD or HD, and MPEG4 is used to create proxy recordings. Proxy recordings are encoded at 1.5 mbps with a 352x288 raster size. This recording is intended for quick low-resolution viewing. Audio on this machine is uncompressed PCM except for MPEG4 proxy recordings. MPEG IMX uses only I frames (File Format section) to simplify editing and will record at 30, 40, or 50 mbps. Third party vendors have also started offering standalone XDCAM recorders.

XDCAM equipment that uses an optical disc as the recording media wraps it in MXF container. XDCAM camcorders that record onto solid-state memory cards use MPEG4 containers for high definition audio/video, and DV-AVI containers for DV video.

The optical disc used in the XDCAM is known as the Professional Disc and is similar to a Blue-Ray disc. Depending on how many optical planes you want to use on the disc at once, from 1 to 4, the disc holds 23—128 GB. This offers a recording time of up to 85 minutes at DV-Cam quality. If four layers are used at once, then the disc can only be recorded on once. It is estimated that the discs, if not used in the write once mode, will have a life span of at least 1000 write cycles and be able to be replayed over a million times. The media's life is projected to be up to 50 years.

Media used with the XDCAM format.

Here is another small, hand-held, storage module from AJA. It has 500GB of storage and spins at 7200 RPM. It works in conjunction with AJA's KiPro record/playback unit, which retails for under $5000.

While DVDs only have a top storage capacity of 5 GB and transfer speeds of 11 mbps, XDCAM dual laser decks have a writing speed of 144Mbps and a reading speed of 170Mbps. This represents a five times' real-time transfer rate for DVCAM material to an external device, and up to two and a half times' real-time transfer for MPEG IMX material.

The optical disc has a cost of media similar to existing professional formats. The blue laser optical recording technology eliminates wear to the disc's surface because there is no contact between the recording laser head and the disc. The laser operates at temperatures of over 1110°F, thus outside temperatures do not affect internal tracking and focus. The optical disc can work in environments much harsher than video tape could.

XDCAM camcorders and decks have almost no moving parts, but the optical block will need replacing periodically; 6000 hours in a deck and 4000 hours in the camcorder. This still rates better than the 2000 hours that tape based scanners would last.

XDCAM camcorders have a three second delay before recording starts as the disc ramps up to speed.

The XDCAM produces hi-resolution MPEG IMX or DVCAM content while generating low resolution proxy file at the same time on the disc. AV proxy content can halve transfer and editing times, and the MXF interleaved format (covered in the File Format section) allows editing to begin instantly, even before the content has finished being transferred.

Each disc has allocated space for 500MB of pure data storage, into which users can put any information they need from Microsoft Office files (Word or Excel files, etc.) to other metadata files.

XDCAM memory cards have also been introduced. The lowest priced cards have a five year life span when used every day to capacity. New memory cards are coming out which will have 1.2 Gbps read/write speed, which will allow it to handle 4K 60p acquisition.

Even today the average replay compartment on a truck carries a combination of video servers and VTRs, as can be seen above the monitors.

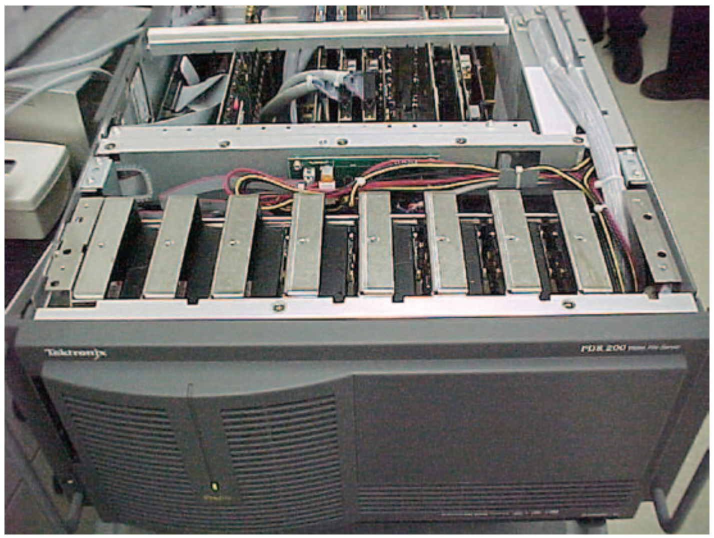
The Video Server

We just looked at systems that have been rapidly evolving from traditional VTRs to become closer to their modern day replacements, video servers. Yet the devices that fit more in the VTR camp operationally than they do in the video server camp still resemble video servers when we inventory the innards of these hybrids. What makes up a video server? First, it's based around one or more processors, internal busses, memory, storage, for both operating and media storage, video processing circuitry, and an operating system that makes it all work as a system.

The XDCAM, and others like it, all have the same systems bundled together. The one difference of any consequence is the media storage, optical storage for the XDCAM, versus hard drive storage for most video servers. So our newer way of managing the storage of media isn't all that radically different from the architectures from the video tape era.

But it should be stressed, where the departure in the two families becomes profound, is in the scaling possibilities with the pure video server. It offers many channels out, much more sophisticated command and control over not only one stand-alone server, but often a fleet of servers, and the possibility of a common pool of storage accessible to all servers at the same time. These attributes will continue to become much more robust and capable as technology evolves.

Today, the limiting factor, as with most technology as we have seen, is bandwidth. Starting with media IO ports onto shared data busses, parallel computing processors, into and out of short term memory, and finally to and from long term storage drives, each step through the subsystems in a server is a series of trade offs for the system designer. At some point, it always comes down to money; spending enough money to pack resources sufficient to outperform the competition, doesn't necessarily mean that customers will sufficiently appreciate the performance boost, and buy more of the product.



An early successful video server that mapped the first steps away from VTRs for many. The photo on the bottom shows the top lid off, exposing the bank of disk drives for video storage.

Like any system you will run across today, be it on a truck, in a factory, or in a bank, it must accomplish its intended task with a viable cost to benefits ratio. So, video servers offer a wide range of capabilities, architectures, and performance, to fit the diverse workflow and budget needs of various situations. A cluster of video servers used for collaborative ingest and editing for news will have a different topology than one used for commercial and program play out, or one used for instant replay and the on-the-fly building of one of more "melds" for a sporting event.

An early widespread server was the Tektronix/Grass Valley Profile. They are still found on some trucks today. As with most computer based systems, the Profile could have a myriad of configurations that could include up to eight JPEG and MPEG decoders/encoders. The storage available to the system could be expanded via SCSI connections to other storage units. The computer control system for the server ran on Windows NT and had its own separate disk drive.

Photo: Grass Valley

Here is the latest video server offered by Grass Valley, the K2, with the front cover off and sitting atop the unit. Like most things technical, it has gotten smaller and much more capable.

Since video servers are, at their core, computers, you would expect them to use computer LANs as a part of their bag of tricks, and they do. Multiple replay servers can be connected together over these LANs, which allow clips recorded on one server to be shared with other servers. For instance, a server that is charged with building a game highlight reel, also referred to as the "meld", can gather clips from all the servers in the truck over the LAN. Another big advantage of the server is that, unlike tape, which requires you to forward or rewind the tape to get to a desired spot (linear operation), servers allow you to go to anywhere in their storage almost instantly. Since you can jump around to different scenes or clips, without having to pass through what was recorded in between, this server operation is referred to as nonlinear.

Besides standard network connectivity, often times video servers will use connectivity geared more towards moving video about, rather than just generic data. One such example is EVS's XNET system. It is a ring that comprises a "round-robin" between all the video servers. Actually XNET uses a hub like most local computer LANs have, where all external devices plug directly into it, but internally it organizes all the connected nodes into a ring.

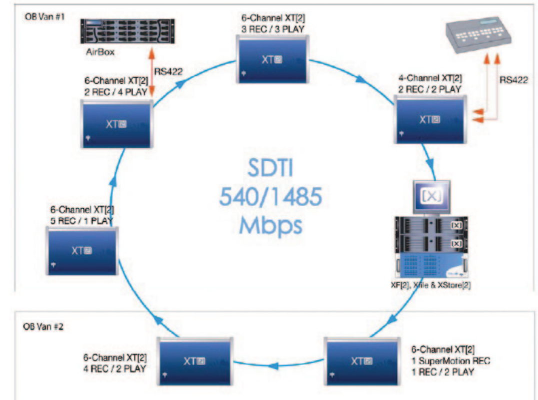


Photo: EVS

Here is the EVS SDTI XNET ring system that interconnects various IO Servers, along with networked storage. The ring can be extended to include multiple trucks at a venue.

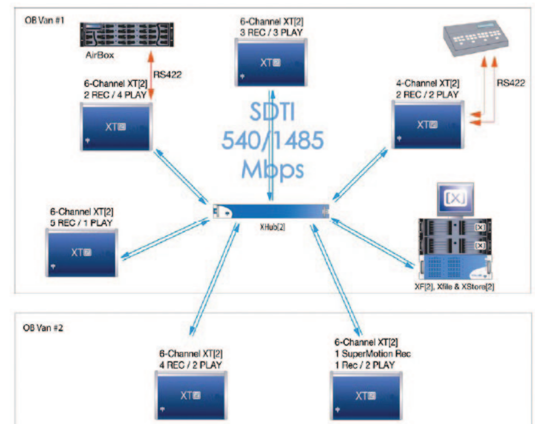


Photo: EVS

In actuality, most systems are organized around a hub, which makes the system look externally like a "star" but in reality preserves the ring topology. This allows for automatic by-passing of nodes that are having connectivity issues, and also allows much more throughput between nodes as it can directly connect two nodes together, isolating them from the others, to give the two communicating nodes all the available SDTI bandwidth.

XNET moves video compressed and stored on one server's hard drive to other servers using digital video hardware. It uses a standard called SDTI (Serial Data Transport Interface). This standard was originally developed to allow HD video to be carried over SD transmission paths via compression of the HD material first.

This allows compressed media on one server to be quickly moved to another. The compressed data is placed in the SDI signal where uncompressed video would normally be. Viewed on any SDI monitoring device, you would see video data at the top of the raster, metadata in the middle and audio data towards the bottom of the raster. It is organized this way, as that was the way it was organized on videotape, which was still the predominant medium when this protocol was developed.

Today, these SDTI transport streams are not based on SDI but now use HD-SDI technology; instead of clock rates of 270 mbps, with about 200 mbps of useable payload, now the clock rates are 1.5 gbps with payloads around 1.2 gbps. These speeds allow faster than real time transfer of compressed data across this network.

So media can be shared via networking techniques, but it can also be shared by setting up storage that multiple servers can use at the same time. There are two prevalent ways that this is accomplished as we will see next; NAS and SANs.

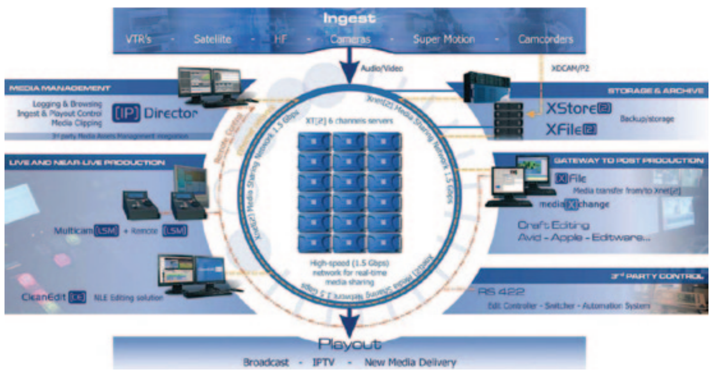


Photo: EVS

High Level Media and control network diagram employed by EVS. Real time media is moved using their SDTI network. SDTI is a standard that uses the HD video framework and puts compressed video/audio data where the video would normally be. Media can also be moved in non-real time using normal Ethernet connectivity. In addition, traditional machine control can be accomplished using RS-422. RS-422 has been in widespread use since the middle 1980s. It is a serial control system that sends a series of bytes representing various machine commands such as play, stop, record, etc. The three common protocols for commands sent over RS-422 are known as Sony Protocol for its inventor, and Video Disk Control Protocol (VDCP), which was developed specifically for use with video servers, as was GV Xten DD-35 Protocol.

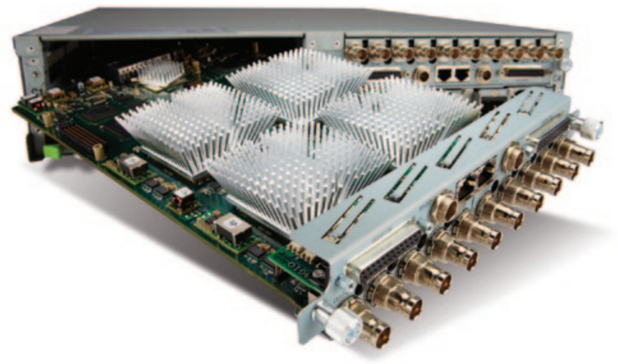


Photo: Grass Valley

Today's servers operate at clock speeds of several Giga Hz. Speed means more heat because the majority of power is consumed during the transition from one digital state to the other. Higher clock cycles mean more transitions per second. The hundreds of tiny columns emulating from the ICs below allow for maximum area to dissipate the heat generated.

Shared Storage

NAS is practically an array of hard disc drives with a network interface. Volumes of HDD NAS are treated by network users as shared network resources. Access to NAS-stored data is provided at the file-level. Accordingly, access permissions for network users are set in the same way as for other shared network resources. In other words, NAS is an advanced file-server that supports multiple network file systems and additional services. Usually, NAS operates on the Ethernet networks using TCP/IP. Thus a NAS is a server that video servers share over a normal computer network.

Among the advantages of this kind of storage is the ease of deployment and maintenance, and also relatively low cost of implementation. One of its serious disadvantages is low performance, as normal Ethernet overhead and other network traffic slows throughput.

SAN architecture differs greatly from that of NAS. The key difference is conveyed even by the names. SAN represents a separate network infrastructure used for information transfer between storages and clients, rather than between a set of networked data storage devices.

Think of a NAS as reaching through a server's networked front door, that is, the IT network that all network devices share. A SAN is a server's backdoor network, where it has a direct path to shared storage that only other devices on a particular SAN can see, but not devices on the main network. A SAN is implemented in a separate network segment. SANs originally were built on SCSI connectivity. At one time that meant big fat parallel cables that looped from hard drive to hard drive and from client to client; in our case that meant from video server to video server. The physical cabling was half of what SCSI was about. It is also a protocol for transferring block-sized chunks of data onto and off of disk storage, and to the proper client or video server. Block-based data access substantially optimizes performance of many resource-intensive applications, including databases, video processing, and so on. Today the protocol is still in use, but it moves over single conductor cables, often fiber. This newer path is known as Fibre Channel.

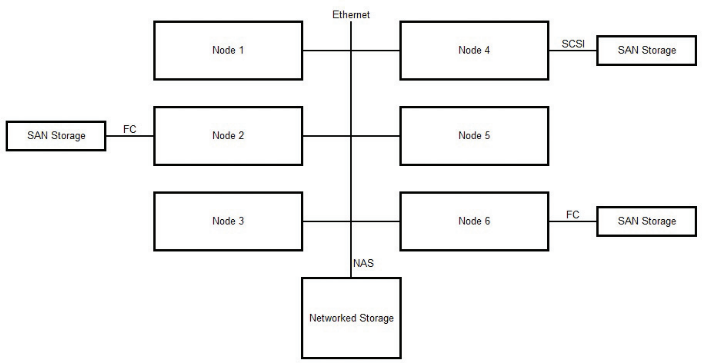
As just mentioned, a SAN operates on data at the block-level rather than the file-level, like a NAS does. The operating system of a SAN client machine treats hard disc volumes on the SAN as local storage, even though it might be sharing that storage with others. Operations on SAN-resources and local disks located physically within the SAN client machine take place in an absolutely identical manner.

A SAN drawback is that backups and mirrors require a block by block copy, even if blocks are empty. A mirror storage group must be equal to or greater in capacity compared to the source volume.

Obviously, the SAN architecture makes heightened requirements to the network throughput. As a rule, Storage Area Networks are based on Fibre Channel (FC) or Gigabit Ethernet equipment. FC protocols, or the NAS network version of SCSI—iSCSI, which run on

Gigabit Ethernet, are used for data transfer. SAN solutions are used when the key factors are reliability, fault-tolerance, and high performance. An advantage to NAS is that devices that use different operating systems, such as Unix/Linux can usually talk to most NAS systems.

A NAS uses Server Message Block (SMB), also known as Common Internet File System (CIFS) which operates as an application-layer network protocol mainly used for providing shared access to files, printers, serial ports, and miscellaneous communications between nodes on a network. It also provides an authenticated inter-process communication mechanism.



This graphic shows how NAS and SAN co-exist and work together. NAS storage is implemented over a LAN, usually Ethernet. The NAS storage is available to all. SAN storage is usually associated with a single node, although to add another layer of complexity a SAN could be shared by multiple nodes.

So what traffic tends to end up on the NAS or the SAN networks? A NAS tends to handle media traffic that is not required to be real time in nature, such as transfer between central storage and editing clusters. SANs tend to handle media that needs to be in real time, such as play out to air. The next question is how are these separate networks tied together? Today that tends to be by a server dedicated to moving data between the NAS network segment and the SAN segment. Remember that NAS storage is usually accessible to many more devices, and devices on the SAN network segments tend to be more exclusive, and limited to devices that actually need fast access to central media. So the server between the two acts as a proxy. Network devices requesting data from the SAN central storage will make their requests to this dedicated server.

Over time, some predict that separate network segments for SANs and NAS will become obsolete as gigabit Ethernets will push speeds high enough that the Ethernet overhead will no longer matter and SCSI protocol over Ethernet, or iSCSI, will win out. Also, NAS has other attributes that SANs don't. Connectivity with a NAS can be via FDDI, or ATM, or HTTP and other protocols.

RAIDS

To safeguard disk storage, usually a number of disks are ganged together, and on each disk, a fraction of the total video and audio data is stored in such a way that if a disk fails, the system can use the rest of the data to fill in the part that was lost. These groups of disks are known as either Redundant Arrays of Independent Disks or Redundant Array of Inexpensive Drives (RAID).

The concept of the RAID was developed at the University of California (Berkeley) in 1987. The RAID can provide protection against a single drive failure, eliminating some or all of the stored data on the array. It does this by generating extra data that is stored and that can be used to determine the data that resided on the failed drive. This usually takes the form of parity data. But the form that the protection scheme can take can be instituted in a number of ways. The various schemes are commonly referred to as RAID levels. The various RAID schemes offer various levels of data bandwidth and protection.

Most RAIDs in use today allow new drives that have replaced failed drives to be rebuilt online. That means that while the RAID is in use, the lost data is re-generated and loaded onto the new drive. This wasn't always the case. Depending on the size of the drive, most can be rebuilt in a couple of hours, offline. Online background rebuilding can take many hours. Some RAIDs can be unforgiving with operational mistakes. On some, if you take a drive offline, the only way to put it back in service is to perform a rebuild of that drive. Also, it should be noted that if for some reason you take a drive offline to pull it from the array, it can take 30 seconds for it to spin completely down and park its heads.

As drives now have spindle rpms above 10K, generated heat is higher then ever. Many RAIDs have backup fans to insure adequate cooling in the face of fan failure. Cooling fans, being mechanical creatures (just like the drives themselves) will eventually fail; it's just a matter of when. Most RAIDs provide over temperature indications, with often more than one warning. One notice as a warning, and a second for automatic shutdown; however, many server vendors override the over temperature shutdown. Drives running above 50 degrees C are undergoing an unwanted stress test.

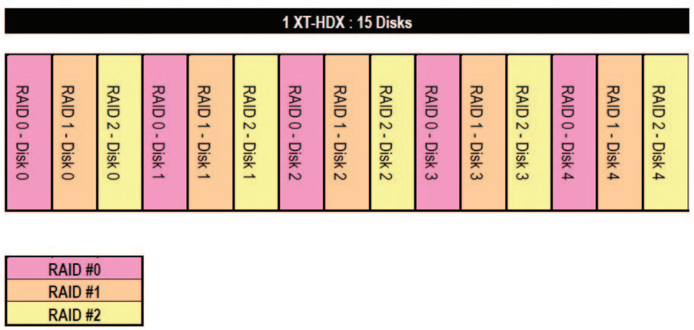
RAID level 0 actually offers no protection at all. What it does do, is provide for high bandwidth by allowing multiple read/writes simultaneously as data is spread to all disks in the array. But if any disk fails, all data in the array is lost.

RAID 1 is simply disk mirroring. All data written to one disk is simultaneously written to a second disk. RAID 1 provides the highest protection, but at the highest cost. This approach quickly becomes too expensive for video servers, but RAID 1 is often used for audio.

RAID levels 2 and higher offer different schemes for using parity data. RAID 2 is not used, as it offered an error correction scheme that has been incorporated into most drives.

RAID Level 3 strips data across all the disks in an array, one byte at a time. There is a drive dedicated to storing ECC (parity). RAID 3 is good for applications that transfer large amounts of sequentially stored data (such as video). This method allows for higher data transfer rates because the data transfer of each disk is available in parallel. Data stripping is what allows for multiple I/O streams out of a single array, as data can be quickly read out and cached, so reads for another stream can commence. Depending on data rates (based on compression ratios) a single RAID is generally able to support around 4 I/O channels. RAID 3 can lose a single disk with no data loss and still provide high bandwidth, even after a disk has failed. RAID 3 is the most common method for protecting video data storage.

RAID 4 and 5 strip entire blocks of data across disks in an array; one block is stored on one disk, the next block on the next disk. This arrangement is good for randomly distributed data such as might be found in a database. Here, access to individual data blocks is much quicker, but large data burst transfer rates are actually lower than RAID 3. RAID 5 shines in applications with frequent transfers of small data files. RAID 4 is the same as RAID 0, as it doesn't provide for data protection, hence it is not used. RAID 5 uses and distributes ECC data over all the disks. RAID 5 provides a higher level of security than RAID 3, but at the cost of system bandwidth. It allows faster read than write operations, because as the data is spread out as blocks instead of bytes, it takes longer to access that data. Many server systems will use RAID 3 to protect video data, and RAID 5 to protect the server's operating system, applications, and database software.



A set of disks need not be organized into a single RAID. They can be separated into multiple sets of RAID storage.

Storage System Topology

The workflows that tend to accompany all this connectivity have become much more extensive over time. In a remote environment, the central thrust is still replays, instant and more compiled "melds". Also there is almost always some pre-production component, which generally has expanded greatly in the last few years. From the event's open, to interstitial features as the game progresses, to a wrap and close, ingest and play out now is often an orchestrated effort.

The producer sets the overall plan as to how the overall recording, storage, and play out capability will be used to convey the desired vision. One or more associate producers usually will handle various elements of that vision. Today it is not unusual to have over 60 input and output channels that make up the throughput capability of the overall storage system. To make it possible for all this capability to result in a cohesive production, there are often multiple operating positions that are control points to affect the actions required to fulfill the desired overall workflow from these systems.

Often, a remote production will need to ingest material from other separate events, or use compilations of footage collected and condensed by the network feed back to the truck. Sometimes this is still "sneaker netted" in; P2 cards and optical disks straight from an ENG camera in the field. Often, a computer, whether full blown server or not, will copy media off these devices and stream them over the NAS or SAN networks for ingest into storage.

The media that is coming into the remote venue from outside sources, and especially the material being recorded locally, will need to be logged. This logging usually results in metadata that now becomes a part of the media and is stored with it.

Photo: EVS

EVS servers. Depending upon configuration, each box can have up to eight IO channels, although boxes that handle 1080p or 3D can have only up to six. Sometimes they are mounted where they are visible, though often they are mounted out of the way, as there are no operator accessible features activated from the box.

Generally, each IO channel in an EVS frame is controlled by an operator replay panel as shown here. Each panel has a RS-422 connection to a port on the EVS frame. An eight channel device would mean that up to eight of these control panels could be connected to a single box. To allow for flexibility, the control cables from each panel go through a patch panel or RS-422 router to allow any panel to connect to any control port on an EVS frame.

All this collected material will need to be able to be located quickly. Many systems record not only high resolution copies, but also produce low resolution versions of the same material, called "proxies". This allows other production folks to be able to quickly view the material and cull segments to be used later in the show, without swamping the network handling the high-resolution material flow. Proxies usually travel over the NAS segments. Often today, separate editing equipment will be on site to ingest from the NAS or SAN segments into editors, to build more elaborate packages than the basic "cuts" only capability that the central storage system has. That said, these central servers are rapidly becoming able to perform their own effects with video.

We must not lose sight of the central mission of these storage systems on a truck, which is replay. The usual architecture is that material coming into the system from various sources goes into servers, which often handle six to eight channels each that are connected to the SAN, into a central or a group of storage systems often over a NAS. The same is true of any servers handling play out duties. There is usually a server, which acts like a bridge between NAS and SAN, which transfers media from the SAN to the NAS and vice-versa. It is a server with lots of storage that often mirrors most of what is in central storage. This server usually acts as the proxy server also. Remember, today a SAN is still generally faster than a NAS. So standalone editors usually receive and send over the NAS into the bridge server for final transfer into the SAN storage for real time play out when desired.

The storage nets of these large systems, whether on the NAS or the SAN side, are expanding out of the truck compound and are often being extended to other venues or even back to the network's NOC. In some situations such as the Olympics, various rights holders and the IBC are on the same storage network, which allows sharing of media, and eliminating the need for multiple organizations having multiple cameras covering the same event.

Another common attribute of servers becoming available is the ability to stream video out of the system either directly to end users on the web, or back through the internet to NOC for processing before continuing out over the web. This practice is still in its infancy, but with the additional viewing angles, highlighting of the action, additional personal and stats-related content, its use will continue to grow.

Streaming implies that a different media must be delivered. With the processing horsepower that can be packed into all types of equipment today, it shouldn't surprise that some servers can convert between SD and HD media internally without requiring any outside help. This is accomplished by using a combination of RISC, FPGA, GPU, and regular CPU devices for real time operation. Besides format conversion, some servers today can even dissolve or perform wipe type effects between two separate channels of video.

Many servers today also come with interfaces to connect with devices such as switchers which allow the server system to appear to be a number of separate record and playback devices under the switcher's control.

Operation

One operator usually controls multiple, often four, channels of video and audio. For football, that means four channels of input and usually another two channels for play out to air. This means that each operator is often in control of a complete server, as many have six channels. Each input is usually fed from a discrete router crosspoint, and each crosspoint might have a dedicated router control panel. During pre-production, the operator loads dozens of clips containing promos, sound bites, animation material and all the elements that allow him to create a diverse range of pieces for use in the live production.

Once game action begins, the operator changes his style of use to that of both a replay operator and a highlight producer. Since each server has four continuous recording inputs, the operator must be continuously aware of what he is ingesting to the hard drives. In addition, the operator has to manage the output channels for efficient coordination of the elements required, such as cueing a promo for the producer to use one minute, and the next minute creating a clip for use later. All the while, the operator must keep track of the overall storylines of the game, so the operator works to tell the story with play lists. Throughout the game, the operator has been creating clips of great plays, reactions and key points that tell the story. With the press of a few buttons, he gathers these clips into a play list and begins the live editing process. These tasks are not completely different from those an editor would perform in an edit suite, but due to the nature of live sports, the game continues to evolve, as do the play lists. An operator may revise a play list several times before it hits air only seconds after the most recent change.

Photo: Corplex

Today, the replay compartment is usually a very busy place, as most cameras in use are recorded into a channel of the video server. There are usually a couple of layers of coordination required to make use of all this potential material. The first is the operator, who besides being a replay operator, must often also take on the role of producer, compiling highlights on his own. The second layer is one or more replay producers whose job it is to let the director know what immediate highlights are available for air, and to filter out what the director should use and not use based on the time available for replay.

For baseball, a replay operator will create clips for each pitch thrown by a pitcher. This enables the operator to instantly create a "pitch-sequence" in a play list. The play list is simply a chain of clips that plays back with transitions. The EVS LSM control panel allows the operator to edit the list live while it is playing back. This "live" editing is what allows the operator to create dynamic highlights that seem to follow the descriptions of the live announcers almost perfectly.

How about car racing? In the traditional world of tape based replays, an operator has two options: recording or playing back. This is limiting, because as everyone knows, in racing anything can happen at any time—and sometimes many things happen in a given moment! This is where the strength of a server comes in; using the system, the replay operator never stops recording! This means action is never missed, and the operator can access it instantly from the hard drives.

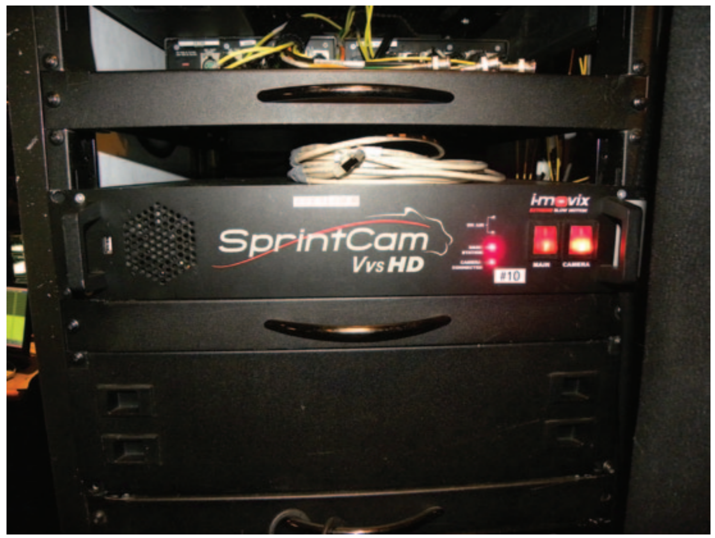
The system also handles high-speed action. In the world of replays, you are normally restricted to slow motion of a 30fps recording, which often looks jerky and is missing critical information. This is due to the frame rate of NTSC television. But, if you speed up the rate at which a camera records an image, you have a high-speed recording that plays back as "Super-Slow Motion." Instead of seeing three frames of a tennis ball hitting the line, the viewer would see nine frames — more detail! These high-speed cameras were traditionally used with a dedicated tape machine recording a specific format, and like the traditional tape machine, were limited to either recording or playback at any given instant. But again, servers can take the Super-Motion to another level. Allowing instant access and continuous recording, server systems make high-speed cameras very beneficial.

Super "Slo-Mo"

The advent of ultra-high-speed cameras has resulted in a change in architecture between the camera's record and playback operations. It used to be that high speed video, up to three times normal was standard, was developed by the camera and sent back to the truck to be recorded and played back. Today, because HD bit-rates are so high, the actual recording and playback of video is all done in the camera itself.

Today, these camera base stations, or the video servers used in conjunction with them, are used to control them during play back. The servers will also record them during play back for future use. The camera itself can also be commanded to retain the clip for future use.

See High Speed Cameras section in the Cameras chapter.



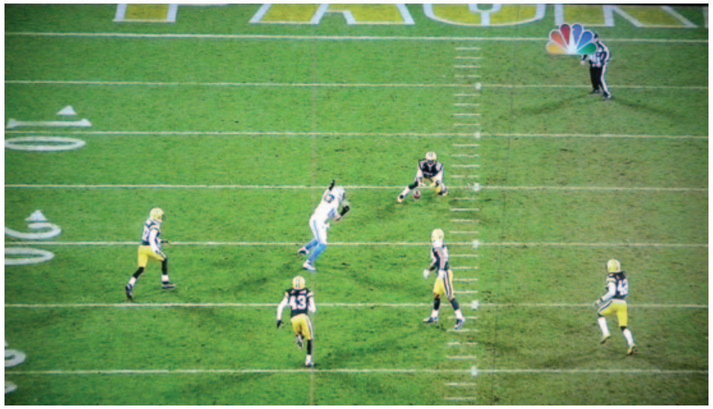
The high speed camera (top photo) works in conjunction with an interface box in the truck. With high speed cameras that can operate at hundreds of frames per second, the camera is no longer relieved of duty during playback. Historically, video from the camera was captured by a VTR or video server, and the camera was done with that video. Now with frames rates, and resulting data rates being so high, the video can't be sent back in real time and is stored in the camera until recalled. The camera then sends back the recorded video at a frame rate no faster than the normal video frame rate. If video is desired from the camera close to real time then the camera drops the amount of frames needed to stay within normal frame rate, which is generally either 30 or 60 frames. EVS and other replay systems have interfaces that are able to control these cameras so they appear as an integral part of the entire replay effort.

4K Replays

While 4K camera video is not sent to your home, or even back to the broadcaster at this point, it is finding increased use at the venue. 4K video is sent from the camera back to the truck as 4 separate HD videos that are stitched back together for use in the truck. One of the uses for 4K video is to combine digital video effects into a server, where the 4K video is recorded and on playback an operator can zoom in on areas of interest. This allows wide areas of the field of play to be covered with a single camera and then provide a zoom for the area of interest during playback.

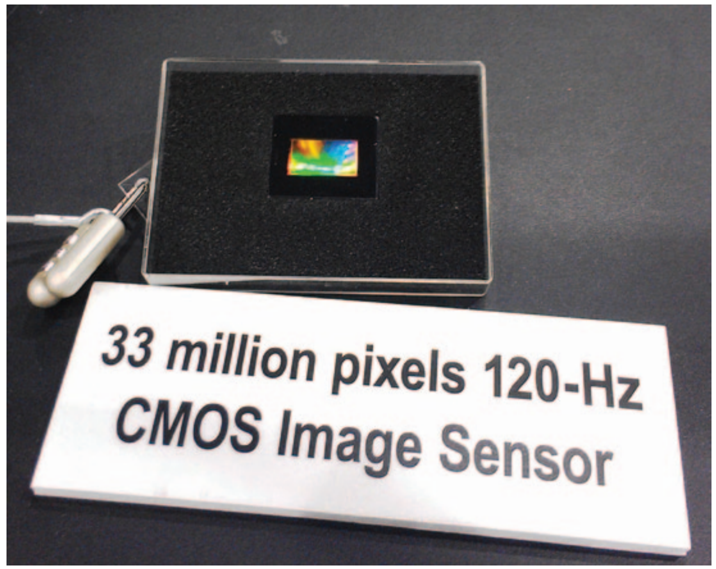
4K camera with one of its two storage modules removed. The camera can operate tethered to the truck or as a stand alone camera that records in 4K. These types of cameras are rapidly finding use in cinematic settings.

Demonstration that shows how a 4K video signal is sent back as 4 separate HD signals and then combined back into one 4K video signal.



Screen shots showing 4K camera in a regular HD broadcast in a replay, and then the replay is stopped and zooms into the area of interest by eliminating most of the 4K pixels down to a point where only the desired number of HD pixels are used.

Here is a demonstration of a system offered by Evertz. On the computer monitor on the right you can see that the operator has highlighted an area of the 4K picture. The monitor on the left shows the resulting "zoomed in" video.



What's next? At NAB 2013 NHK demonstrated an 8K camera. This camera is larger than regular HD, and 4K cameras, and due to its operating speed, understandably produces a fair amount of heat.

Conclusion

Very crude attempts at replay were started in the early 1960s using the VTRs of the era. But these machines were not up to the task, and at best, if they could be rewound and put into play, and if they locked in time real time only playback could be attempted. Cameras were generally not isolated, there were generally only a few at most events, and the on-air feed was the playback feed. In the mid-1960s when Ampex introduced its analog disk based recorder, the replay could be nearly instant and in slow motion, or up to 2x play for that matter. It was revolutionary, but due to cost, rare, being used only on major network broadcasts. The advent of the one-inch VTR made replays ubiquitous and they were used on most events. But even with one-inch VTRs on the scene, there still weren't that many simultaneous replays available. Building highlight reels generally took time away from the immediate replay effort and could often only be performed during halftime intermissions. It wasn't until the adoption of video servers where storage became centralized and available to many at once, that the review of what had happened, became the mainstay of many events.


C - BRUCE
[C - IN]
[GREEN - IN]
GREEN - BRUCE
X - ERIC
[X - IN]
[Y - IN]
Y - ERIC
[A - IN]
[B - IN]
A - ISABEL
B - ISABEL
PGM