It's Called Storage Today

Figure 5.22 This HD truck's tape compartment includes Sony HDCam high-definition VTRs, with a standard-definition Sony DVW-A500 in the center. The HDCam subsamples video at 3:1:1 at eight bits, so as not to have to compress the resulting data as much. The data rate onto tape is 140Mb/s. (Source: Tim Stoffel)

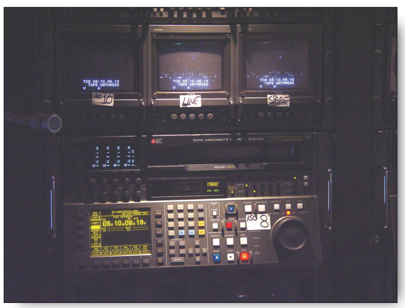


Figure 5.21 A mainstay on trucks today is the Sony DVW-A500 Digital Betacam. The "A" in the designation means that this digital VTR can play back analog Betacam tapes, even though the two formats are physically quite different. It's able to do this through DT head and TBC magic. The Digital Betacam uses a mild 2:1 JPEG compression. It records 10-bit component video and 20-bit AES audio sampled at 48KHz. The bit rate onto tape is 127.8Mb/s. The evolution from tape to server now taking place in the tape compartments of most production trucks is also happening inside the VTR. Sony introduced the hybrid VTR with their SX series. A hybrid VTR is a VTR with an internal disk drive. This would allow one VTR to edit or assemble raw footage onto its internal disk drive by itself, without using a second VTR.

Tape's Going Away, Just Not Yet

Common wisdom in television's technical circles is that videotape, and the contraptions that play those tapes, are on their way out. Some people in the mid-90s predicted a tapeless world by now. Problem is, storage on tape is still cheaper than storage on disk, although that gap is narrowing. The other good thing about tape is that, short of a fire, it's hard to lose your whole tape library. If all your video is on disk storage, though, and something happens to that storage, well…. And that has happened. 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 redundant arrays of independent disks, or RAID. The RAID is connected to a specialized computer known as a video server. As we will see, if the risks are managed correctly, the video server can be a much more powerful device than a number of VTRs combined.

The strange thing about the demise of videotape is this: We had one tape format at first, then a few in the '60s, '70s and '80s, and today there are more than a dozen different tape formats. A large part of the reason for all the formats is that almost all formats today are digital. That means that audio and video are recorded onto tape as data. As we have mentioned before, digital data requires lots of bandwidth, and compression is used to limit that bandwidth to make reasonably priced VTRs ($20,000 to $60,000 is considered "reasonable" for broadcast). Different manufacturers have differing philosophies on video and audio compression and what size tape is best, hence all the formats.

Cameras, and especially VTRs, have seen a reversal in

Figure 5.23 VTR sophistication has also created smaller machines, allowing editing to be done anywhere, including in the car or the cab of the Ku truck. The portable editor here contains two VTRs in one package. (Source: Sandorse)

Figure 5.24 Right: Today's tape compartments are usually a combination of analog and digital VTRs, along with an ever-increasing tally of video servers. A couple of operators often control two or even four VTRs in record and replay playback. A replay producer, whose shoulder we are looking over in the bottom right photo, often coordinates replays. Some productions require so much replay capability that an entire trailer is devoted to VTRs, as is the case with the truck at lower left, which is used for ABC's Monday Night Football. It is shown here at the Indianapolis 500.

technology flow. It used to be that the RCAs, Sonys, and Panasonics would develop technology for their high-end product offerings, such as the ones for broadcast, and that technology would trickle down into consumer gear. Now the opposite is becoming 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. The Beta tape morphed into the high-end Sony Betacam VTR family. VHS did the same thing, becoming Panasonic's MII. The tape and some of the tape transport technology is the same, but the electronics are definitely high-end as compared to consumer gear.

As with other kinds of equipment, the truck industry has settled on a few VTR formats. Over the years these have migrated from analog Ampex and Sony Type C one-inch VTRs in the early 1980s, to Sony Betacam in the late '80s through the '90s. Many still abound today: Sony Digital Betacams and Betacam SX, a few HDCams, and some Panasonic D5s.

The Video Server

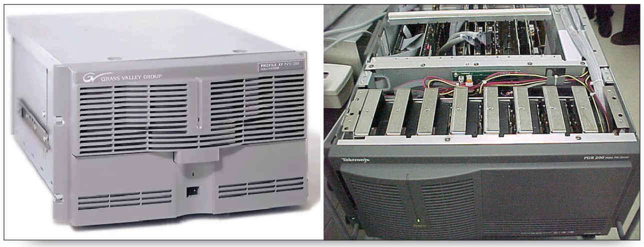


Figure 5.25 *A couple of related video servers. The left one is from a recent offering from the Grass Valley Group, the right one is an earlier member of the family when Grass Valley and Tektronix were related corporately. The exposed top on the older version reveals the eight disk drives at the front that make up its RAID. RAIDs can be internal to the server, or located in a separate rack-mounted assembly (used especially when storage requirements are large). The boards behind the RAID are a specialized computer running on Windows NT that controls the electronics that actually process audio and video.*

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 can gather clips from all the servers in the truck over the LAN. Another big advantage of the server is that, unlike tape, which require 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, server operation is referred to as nonlinear.

Super "Slo-Mo"

The demand for increasingly interesting replays has actually initiated a couple of developments in the camera domain. One is a camera that captures more frames per second than normal cameras. The argument can be made that high definition is not just more information, or pixels, per frame of video. In some cases, presenting more pictures a second can provide a better view of some activities. Television standards around the world generally have frame rates based on the local power frequency because in television's infancy, when electronic technology was also much less advanced than today, a television receiver would be much more reliable if it received a signal that was in sync with power. In North America the power frequency is 60Hz and in much of Europe it's 50Hz, so these two areas of the world have 60 and 50 television field rates. That means that the television receiver or monitor scans 60 or 50 times vertically every second. Again, because of technological and bandwidth limitations set early on, only half the horizontal lines making up the picture were sent with each vertical scan. The rest of the picture's horizontal lines were sent the next field. One field had all the odd lines, and the next field laid the even lines in

Video Server Operation, by James Stellpflug

Figure 5.26 Controller for multichannel server. One operator can control many channels of video and audio.

Let's start with a football game, using a six-channel EVS LSM (Live Slow Motion), which uses a 4CAM system to record four cameras into the server. The system has two outputs for air channels. Each input is fed from a discrete router crosspoint, and each crosspoint has a dedicated router head. During pre-production, the operator loads dozens of clips containing promos, sound bites, animation material and all the elements that allow the LSM to create a diverse range of pieces for use in the live production.

between the odd. This is known as interlace scanning. The bottom line is that it takes two fields to send one complete television frame. So the frame rate is only half 60, or 30 frames per second (Europe is 50 divided by 2, or 25).

This means that the swing of a bat or a golf club can only be broken down into 30 snapshots per second in a fast-action replay. The super "slo-mo" camera, pioneered by Sony, output video at 90 frames a second. Originally it had a VTR that recorded at the same 90-frame rate. Now the camera often feeds a server that records at the same high frame rate. Concurrently the camera outputs a standard 30-frame signal, so that it can be used live also. On playback the temporal resolution, or amount of new information presented over a given time, is greater. The only drawback from an operational standpoint is that the fastest the replay can be played back is at 1/3 speed, as that is equal to the normal video frame rate, and since camera and recorder are operating at higher frequencies heat buildup is greater.

Figure 5.28 *Replays are currently used by the NFL in reviewing plays disputed by coaches or replay officials during the last two minutes of each half. The NFL has its own server system that takes live action and replays from the production trucks. Four people in a booth operate the NFL review system, which is generally on the press box level. One person with an officiating background is the only one in contact with the referee on the field. This person will let the field crew know when there is a coach's challenge, which is indicated by a buzzer system that each coach has. This same person will, in the last two minutes of each half, decide when to initiate a officials review, and decide what replay views are looked at by the referee in all reviews. The referee has 90 seconds from the time he dons the headsets to make a ruling. Another person will initiate record of live action and replays coming from the truck. Another operates the slow-motion playback under direction of the person talking to the referee. A fourth person will continually announce to the other three the status of activity between plays, so they know when the players are in the hurdle, at the line and when the next play has commenced. This means that the time for any further review of the previous play is over.*

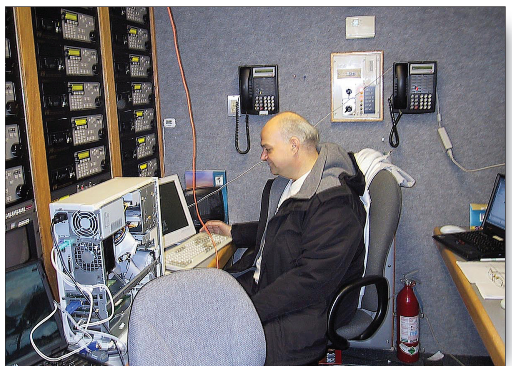
Once game action begins, the operator changes his style of use to that of both a replay operator and a highlight producer. Since the LSM has four continuous recording inputs, the operator must be continuously aware of what he is ingesting to the hard drives. In addition, the operator has to manage the use of the output channels for efficient coordination of the elements required during the show, 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, trained to see what the overall storylines of the game are, the operator works to tell the story with playlists. 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 playlist and begins the live editing process. These tasks are not completely different than 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 playlists. An operator may revise a playlist several times before it hits air only seconds after the most recent change.

In baseball for example, an LSM operator will create clips for each pitch thrown by a pitcher. This enables the operator to instantly create a "pitch-sequence" in a playlist. The playlist is simply a chain of clips that plays back with transitions. The LSM 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.

Figure 5.27 The GUI that a server operator uses to manage clips that have been saved and will be used later, either separately or as part of a highlight package.

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 minute and sometimes many things happen at any given minute. This is where the server comes in. Using an LSM system, the replay operator never stops recording! This means action is never missed, and the operator can access action 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, the Super-LSM systems make high-speed cameras very beneficial.



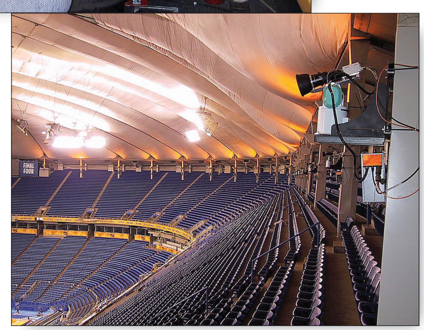


Figure 5.29 *Eyevision is an extremely interesting twist, literally, on the instant replay. This is a system of 30 or more cameras spaced 80 feet apart that have a 270-degree view of the action. One of the cameras is designated the master camera. Through a computer and robotic system developed through the joint efforts of CBS and Carnegie Mellon, that is now being refined with additional partners Core Digital Technologies and Princeton Video Imaging, all the cameras are made to have the same lens focal length and centered on the same action as the main camera via robotic pan/tilt control. The only difference is that each camera has its own unique angle of view. All cameras are recorded on a separate VTR, or channel of a server, and then played back in concert. At an interesting point in the replay, action can be stopped and the output of the camera closest to the current on-air camera's playback can be selected, then the next closest, etc. in sequence. This gives the illusion of swinging around from one view to another. Kind of like a Skycam move, only after the fact. Eyevision's ability to step around from camera playback to camera playback to change viewing angles might only be the start. With 3-D geometric mapping of what all the cameras see from their various angles, some heavy video processing might enable a couple of virtual reality features to fall out. One would be to "erase" players in the scene that have no bearing on the play. The second item would be to use the collected picture of 3-D space from all the camera angles to create angles of view that don't physically exist, such as the view from the backfield in football, or the viewpoint of the pitcher in baseball.*

From instant replay to instant virtual reality, and from discrete sources to one composite omnidirectional source, all because someone way back when thought it would be nice to get a second look at the action. ■