SECTION ONE: THE DIGITAL VIDEO PROCESS

The simplest definition of digital video is the representation or encoding of an analog video signal in digital bits for storage, transmission and display. If you have access to the World Wide Web, chances are you have viewed digital video files. Examples include CNN news clips, movie trailers and, of course, the popular dancing baby! Digital video files pervade the analog television and 70 mm film world. Most special effects, such as Godzilla trampling a building or a polar bear sipping a soft drink in an advertisement, are created by editing digital video files. If you have rented movies on demand in your hotel room, played a DVD, a video game, or used a Direct TV satellite dish; you have experienced digital video through your television set.

Digital video is a growing presence in the academic arena, from digitized course lectures to archival footage housed in the campus library. Video conferencing--for collaboration, Internet-based communication and teaching--is an important digital video service. Video conferencing is addressed by ViDe in the Video conferencing cookbook. Digital video on demand, another key service, is defined for this white paper as the creation, storage, transmission and display of archived digital video files in an IP network environment. Digital video may be streamed to a computer, so that playback begins before the entire video file is received, or non-streamed, requiring that the entire file be downloaded be playing. Streaming videos may be served as multicast or unicast streams. Video on Demand generally refers to unicast, where a single video file is requested by a user and streamed to the user's computer for playback. Multicast--the transmission of a single digital video file to multiple users in a scheduled environment, is included in this digital video on demand white paper for convenience, since most vendors providing on-demand video files to a single user (unicast) also provide products for multicast of both stored files and live broadcasts.

There are three basic components in the digital video process.

Step One: Encoding or Digitizing the Video

The analog signal, which can be a direct broadcast or a videotape, such as a VHS cassette, must be digitized. An encoder card accepts an analog signal through a cable into an interface card and feeds the signal into the encoding hardware and software to encode the video into digital form.

Encoding is a simple concept: the video analog signal is encoded, or represented, in digital bits that can be read and operated upon by a computer processor. All digital files--whether a textual document, an image, a program, or a video--are representations of information in bits.

One meaningful element in graphical digital media is the pixel, or picture element, which is a two-dimensional base unit of programmable color, also represented in bits. A pixel is a logical element, so that its representation can vary based on factors such as bit depth and screen resolution. The color expressed by a pixel is a blend of some component of the red, green or blue color spectrum. The human eye has photoreceptor cone cells that respond to color in the three spectra, so three mathematical representations of color--red, green, and blue--are all that are needed for digital color reproduction. Like all other digital information, colors are represented by bits. More bits (8-bit, 16-bit, 24-bit, etc.) allow more precise definition of the exact blend or hue from the red, green, blue color spectrum. Analog colors are translated to pixels in the RGB digital color space. Digital video uses a non-linear variation of RGB called YCbCr, where Y represents luminance, or brightness, and CbCr represents chrominance (chromaticity), or "pure" color, in the absence of brightness.

The number of pixels displayed on your computer screen, along the horizontal axis and the vertical axis, is defined as the spatial resolution. Broadcast-quality digital video (CCIR 601) is commonly displayed with 720 x 480 resolution.

Video is more than color, however. Video requires multiple frames showing redundant information and fractional changes to create the illusion of motion across time and space. At some point in your childhood, you probably duplicated the illusion of motion by drawing stick figures on cards and flipping them to create your own "cartoon." The more redundancy between frames, the smaller the change from one frame to the next, the smoother and more continuous the illusion of motion on your television or movie screen. Video encoding algorithms take advantage of this redundancy to compress video, encoding only the difference between frames. This process is known as temporal compression. The decoder at the client end stores the information that does not change from frame to frame in the buffer to refresh the displayed frame as needed.

To convert analog video to digital, each frame must be digitized using encoding hardware and software. Encoding systems can be as inexpensive as a $300 card fitting into an available slot on a multipurpose microcomputer to a $10,000+ stand-alone system, which requires a dedicated microcomputer. A good tape deck and analog monitor are also usually required for the encoding process, depending on the requirements of the selected video card/encoding system.

A video-encoding card accepts analog inputs from a VCR or a video camera and converts the analog format into a digital video file. Encoding hardware and software vary greatly in cost and therefore support a wide range of functionalities, including, as the cost increases, higher quality output, separate input for video and audio, faster encoding, multiple file and batch file processing, analog output (e.g. digital video back to analog videotape), uncompressed conversion and, at the present time, a range of encoding formats, including M-JPEG, MPEG-1, editable MPEG, MPEG-2, Video for Windows/ActiveMovie, and QuickTime.

Video cards are available for proprietary formats such as Intel's Indeo®. Video cards create digital files that can be opened by editing software, such as Adobe Premiere®. Video editing packages allow you to make changes to digital video, such as adding credits or special effects, cutting or adding frames, merging digital video clips, and outputting the created movie to a range of digital file formats, for playback in a variety of ways, through the use of incorporated software or plug-ins.

When digitizing video, each frame must be converted to digital form. In addition, the audio track accompanying the digital video must be converted and synchronized to the video for playback. A straight digitization with no compression requires more bandwidth and processing power than desktop computers can handle. Currently, compression must be employed to convert analog audio and video so that it is usable at the desktop. Compression reduces redundant information so that meaning is not lost but file sizes are reduced to manageable form. When you "rent" a movie in a hotel room, you are seeing a slightly-compressed (MPEG-2) version of a movie you might have seen in uncompressed analog form a few months earlier at a movie theater.

Uncompressed digital video can be created on high-end platforms and broadcast in real time or stored for later use. The Society of Motion Picture and Television Engineers (SMPTE) and the Institute of Electrical and Electronics Engineers (IEEE) develop and manage standards for uncompressed digital video. These standards include CCIR-601 for television broadcast digital video, in the resolution of PAL, NTSC, and SECAM; SMPTE 259M for the transport of CCIR-601; and SMPTE 292M for the transport of high definition television (HDTV).

The Moving Picture Experts Groups, known collectively as MPEG, are responsible for developing and maintaining digital video and audio encoding standards to address a wide range of commercial and educational needs. MPEG employs established procedures for the development, adoption, testing and review of digital multimedia standards. Standards are published and made freely available to commercial developers, although reasonable costs for some technologies may apply. MPEG standards are international standards that insure video encoding systems will create standardized files that can be opened and played at any desktop with a standards-compliant decoder. MPEG encoding standards are discussed in more detail in Section 2.

In the past few years, digital video cameras have become available, in commercial and consumer-quality models. A high performance serial bus, IEEE P1394, popularly known as FireWire, was developed by Apple Computer but now supported by many vendors to support data transfer rates of 100, 200 or 400 Mbps.

These high transfer rates mean that digital video can be transported directly from the digital source (camera, DVD, etc.) into the microcomputer with no processing delays. FireWire streams video data off a hard drive in real time without computer assistance. FireWire supports up to 63 devices on a single bus, which can be connected in a star, tree, or daisy chain pattern, and allows 1,023 buses to be bridged together. Addressing is dynamic and allows devices to be connected without rebooting the computer.

FireWire transfer speeds, currently at 100-400 Mbps, will increase to 800 Mbps/multi-Gbps in the next release--1394B. The high transport speeds can result in latency problems, requiring significant buffering capacity, as can a heavily-loaded PCI bus, but these problems will abate as FireWire integration becomes the norm, and microcomputers are designed for FireWire integration.

One issue with digital video camera content creation is that the resulting digital video files are generally in formats not currently supported by digital video client/server systems. When investigating digital video cameras, insure that a method exists to output the file to a standard format (AVI-to-MPEG or directly to MPEG)--whether through an editing program such as Adobe Premiere or a transcoding system such as Heuris.

Step Two: Sending Digital Video to the Desktop

Once a video is created, it is stored and then transported to the desktop for playback. Digital video created on a computer can be stored on the computer, opened and played back, just as a document is opened in a word processing program for reading, editing and printing.

A server must generally be employed to store and share a video over a network--whether a campus or building LAN or the Internet. Digital service includes real time broadcast, non-streamed downloading or streaming to the desktop. Video service may be multicast ("one to many") where one video stream is served to many viewing clients or unicast ("one to one") where one video stream is served to one viewing client.

Real time broadcasting converts analog video to digital on the fly. Analog video is received by the video server directly from a broadcast feed or a video camera, encoded in real time, and then served as a multicast video stream to many clients. Real time broadcasting also includes the real time delivery of files already in a digital format, such as a digital camera or satellite transmission.

Video files are meaningful only when forward progression, providing continuity of information, is maintained. A cartoon coyote cannot be running off the cliff in one frame and standing on the edge of the cliff, looking down, in the next frame, if the video file is to make sense to the viewer. The coyote also cannot be running swiftly to the cliff in one frame and moving slowly and jerkily in the next. Video data must be played in the correct order, with little or no packet loss, and with smooth, continuous timing or else essential information will be missing. To insure that video files are usable at the desktop, the frames must be received in order and timed for playback. Digital video can be received at the desktop for playback in two ways: non-streaming or streaming video.

Non-streaming video requires that an entire video file be downloaded and lacks the timing functionality for smooth packet streaming. A video server is not required to store and serve non-streaming digital video.

Any server can store and serve non-streaming video, or the non-streamed file may be stored on the microcomputer hard drive or a CD-ROM and played back.

Digital video functionality, for opening the non-streaming file and playing it back on the client machine, is provided by the client software. Downloaded video is an option when the latency (elapsed time) required for the download process, which can range from several minutes to more than an hour, is not an issue.

Non-streaming video is also employed when a video server is not available to provide streaming. Non-streaming video files may also be provided when the maximum number of concurrent video streams supported by a video server has been exceeded. Most video server vendors do not support download of non-streaming videos. If an institution wants to offer nonstreamed video files for download, to insure high availability for the files, an FTP server or other download site must be separately provided.

Streaming video begins playback on the client as soon as enough of the video has loaded to begin and sustain playback at a continuous rate. Cache is established from random access memory (RAM) on the client desktop and is used to receive the file, insure that frames are in the correct order, establish timing, refresh compressed frames and check for dropped packets. The video file continues to download into the client cache even as the beginning of the video is being viewed. Video streaming relies on technology at the video server and at the client, such as caching and control bits, to receive and assemble a video in which all data bits play smoothly, in progressive frame order.

Video streamed via the Web must be transported within the IP architecture. Streamed video has low tolerance for the enforced reliability of TCP, which would keep an application waiting for the retransmission of dropped packets. UDP (User Datagram Protocol) is frequently used in place of TCP as a transport protocol for real time applications, such as digital video.

UDP uses the Internet Protocol (IP) to transport a data unit ("datagram"). UDP supports digital video because it does not divide the data stream into packets for reassembly at the client end. However, UDP also does not order the datagrams into the correct sequence. Applications using UDP must insure, at the receiving end, that the complete message has arrived, in the correct sequence order.

RTP (Real-time Transport Protocol) is a UDP protocol that provides payload type identification, sequence numbering and time stamping. RTP allows for packets to be transported out of order and reassembled in correct order at the receiving end. Digital video has low tolerance for disordered packets and dropped frames. It is used on the MBONE, for interactive audio and video, particularly conferencing sessions. RTP is used with a companion protocol, RTCP (Real-time Control Protocol), which provides periodic control packets to an application to monitor the quality of the data distribution.

RTSP (Real-time Streaming Protocol) is an application-level rather than a simple protocol, since it works with many transport protocols--TCP, UDP, RTP, and IP Multicast. RTSP was designed to support streaming multimedia in unicast and multicast applications. It provides increased functionality at the client end for playback, seeking, etc. and has been described as a "video remote control" for the computer. Among other features, RTSP allows for interoperability between server and client implementations from different vendors. RTSP can be used with RSVP to establish and manage reserved-bandwidth streaming sessions. Progressive Networks' Real Player G2 is an example of an RTSP client.

RSVP (Resource Reservation Protocol) provides Quality of Service (QoS) by allowing an application invoking RSVP to reserve end-to-end bandwidth, memory and CPU resources sufficient for the demands of the application. RSVP requires that all network components work together to provide guaranteed resources for the application, so all components--hosts, routers, hubs, etc.--must support RSVP. Although RSVP is a fairly mature standard, it is not heavily implemented, due at least in part to the requirement that all network components support the protocol.

IP Multicast supports one-to-many service for a data stream. All routers in the network infrastructure must be IP multicast-enabled. Some multicast applications, such as Progressive Networks' RealSystem G2, are able to bridge non-multicast-enabled network segments. A process asks its host for permission to join or leave a group. IP multicast-enabled routers query their groups to identify the processes currently belonging to each group. On the Internet, IP Multicast is implemented on the MBONE and, increasingly, by service providers who will multicast your video to your authenticated users for a fee. Multicast is supported natively on advanced networks, such as vBNS and Abilene.

Quality of Service (QoS) provides a mechanism whereby a client can request priority access, sufficient bandwidth and other network service characteristics to guarantee acceptable application performance. QoS sounds deceptively simple but is difficult to implement since the QoS protocol (such as RSVP) must be enabled on all network devices to insure that the bandwidth allocation is supported across the network. QoS becomes increasingly important as more clients request unicast video applications or participate in a multicast transmission. Most networks currently utilize admission control (users beyond a prescribed stream limit are denied service until a stream is freed) or "best effort," where all users share bandwidth equally and suffer equally, as bandwidth utilization approaches capacity.

Step three: Playback at the Desktop

When a streamed video file is received at the desktop, the file type must be recognized, through information provided in the header, and then opened. The file must be cached until sufficient data is received to allow for smooth, continuous playback. Playback includes controls such as forward, reverse, stop and play, as well as freeze-frame, content bookmarking, audio volume control and sizing of the viewing window. Client software, frequently a web-browser plug-in or browser helper-application, provides this functionality. If you visit a web site with audio or video, you are usually alerted to the plug-in viewer needed and provided a link for free download of the plug-in. RealVideo, Windows ActiveMovie (AKA NetShow), and QuickTime are common file formats with freely-available viewers.

In the past, different client players were required for different file types. Currently, many standard client players have extended their client capabilities to recognize multiple standard file types, such as MPEG-1 (.mpg), ActiveMovie/Video for Windows (.avi), QuickTime (.mov) and RealVideo (.rm and .ram). Progressive Networks' G2 Real Player, with plug-in extensions, and Microsoft's Windows Media Player can open and playback multiple digital video file formats, for example.

Additional functionality provided by the client, in collaboration with the server, can include permission to save a file to hard drive, optimization of playback based on network connection (e.g. 28.8 KB, 56 KB, T1, etc.) and auto-selection of video file format for a video asset transcoded in multiple formats, usually based on the bandwidth capability of the client network connection.