

Pro-AV: Inside Media Streaming

SOURCE: October 2004 issue· POSTED: 05/12/05

The technology behind what's quickly become a lucrative mainstream media tool, and what you can do to grab market share now.

Streaming media is everywhere. As of July 2004, 51 percent of all home Internet users had a broadband connection. This increased bandwidth enables streaming media to enter the home much more easily than with slower connections, whether it's in the form of an Internet radio station, a movie preview, or the local weather forecast. With the Internet acting as the driving force behind it, media streaming has spread to many different markets close to most system integrators' hearts, including corporate, retail, education, and judicial.

The approval of H.264/MPEG-4 part 10 (AVC) has cut bandwidth requirements for quality video in half, increasing the applications of this technology to the point where it's now considered mainstream. In fact, some experts estimate that 100 percent of Fortune 1,000 companies are already using it — that's not even counting the nearly ubiquitous applications in entertainment companies, cable head-end providers, broadcasters, and government institutions. To understand how media streaming works, let's take a look at the technology behind it.

Standards

To create a stream, the content (video, audio, pictures, etc.) must be transformed into a digital format. This happens through the use of a codec, which uses mathematical algorithms to transform the content into the digital realm and compresses the file size in most applications. Media streaming codecs fall into one of two categories: standards-based and proprietary.

Standards-based codecs follow an open source guide to compliancy developed by standards organizations such as the Moving Picture Experts Group (MPEG) or International Telecommunications Union (ITU). These standards organizations are comprised of numerous companies that hold an interest in the development of the open standards. The reason for having an open standard is to increase interoperability in the marketplace, allowing products from different manufacturers to communicate effectively.

Proprietary codecs, on the other hand, are most commonly developed by a single manufacturer to fit a specific need. Examples of proprietary codecs are VC-9 (The Windows Media 9 video codec) and RealVideo. Performance of proprietary codecs tends to be of higher quality because developers only need to worry about compliance with themselves and not the entire industry. But there is a tradeoff for this increased quality: You lose compatibility with products that don't employ the same codec.

Because proprietary codec developers like to keep the behind-the-scenes information a well-kept secret (after all, that's how they make money), let's focus on the standards-based MPEG formats for describing codec technology.

Formats

In the beginning, there was MPEG-1. Becoming official in 1992, this standard was the first attempt from MPEG to deal with digital video and the problems with playing video files on computers and other playback devices.

In November of 1994, MPEG-2 was approved, which provided a mechanism for providing higher quality video with increased compression rates. MPEG-2 is the most commonly used standard for entertainment video today, including DVDs, digital cable, and digital satellite TV. Advanced Audio Coding (AAC) file types, recently made popular by the iPod and iTunes, are part of the MPEG-2 standard. MPEG-3 never came to fruition as it was slated to be the HDTV standard. Instead, it became a part of MPEG-2. MP3 files, the ones that revolutionized the digital music world, are actually part of the MPEG-1 standard's layer three audio compression.

The first parts of MPEG-4 were approved in 1998 with a focus on low-bandwidth transmission of audio and video for videoconferencing. Since its original inception, MPEG-4 has morphed into something much larger, involving interactivity and a digital rights management scheme called Intellectual Property Management and Protection (IPMP).

MPEG-7, also known as the Multimedia Content Description Interface, is all about the metadata that describes the content, quality, condition, and other characteristics of data. This standard goes beyond the simplistic storing of title, artist, and album to include descriptions of individual elements, additional cataloging divisions, and technical statistics.

The latest and most encompassing standard is MPEG-21, which will act as the framework for the delivery and utilization of multimedia content. Viewed as the all-encompassing standard, MPEG-21 gives content distributors complete control over the entire media delivery system.

The surge in high-definition video has also prompted a war among codecs to determine which will be the next publicly accepted standard (see "Two's a Crowd" on page 64 of the April 2004 issue of Pro AV).

With the two successors to the DVD format (HD-DVD being developed by the DVD Forum and Blu-ray developed by 11 consumer electronics manufacturers) nearing their final specifications, the different codec developers are jockeying for a spot to be a mandatory part of compliant devices. The DVD Forum recently took the VC-9 codec off pending status, and it will now be a part of its final specification along with MPEG-2 and MPEG-4. The Blu-ray specification already includes the MPEG-2 codec, and its creators are considering including VC-9 and the new MPEG-4 AVC High Profile (formerly AVC Fidelity Range Extension) codec approved in July 2004. Because developers are constantly tweaking codec technology to get the most out of bandwidth, it's difficult to declare a clear winner.

How MPEG video coding works

The basic idea behind a compression algorithm is to keep only the information that is really needed (entropy), and throw away the rest (redundancy) during the encoding process. Then, during the decoding process, the algorithm tries to reconstruct what was thrown away during the encoding process using what information is available. The more efficient a codec algorithm is at removing the unnecessary parts, the more it equates into lower transmission bandwidth and smaller file size. MPEG's compression algorithm works by only keeping track of changes from frame to frame. For example, if you have a person standing in front of a black background, the only parts of the image that change from frame to frame are those occupied by the person. The parts of the frame that are black are considered the redundant information and are not necessarily transmitted for each and every frame.

An MPEG video sequence consists of three different picture or frame types that form a Group of Pictures (GOP). A GOP sequence starts with an "I picture" that acts as the reference on which the rest of the group is built. The I picture is also referred to as the intra-coded picture because the compression algorithm applied to this frame occurs only within itself and doesn't rely on surrounding frames. This process is similar to the compression of a JPEG image. The remaining pictures in a GOP are forward predicted (or P) pictures and bi-directional (or B) pictures. The first P picture in the GOP uses the I picture as a basis when decoded, utilizing both the addition of difference data and motion compensation to regenerate the parts dropped during encoding. Each subsequent P picture uses the previous P picture as its reference until a new I picture is transmitted starting the next GOP. The bi-directional B pictures are the filler in the middle and rely heavily on the surrounding I and P pictures.

These pictures are decoded using difference data and motion vectors from the I and P pictures immediately preceding or following the B picture in the GOP. Because the B pictures can be backward predicted from a P picture that follows it, the pictures are sent out of order with the P pictures preceding the subsequent B pictures The final ratio of I to P to B pictures is adjustable to the stream's bandwidth and quality needs. An increase in the number of B pictures will reduce the required bandwidth, although quality will suffer as these are the most highly compressed of the three. Increasing the number of I pictures will increase the quality. However, because the I pictures are the largest in terms of file size, the overall compression on the stream will decrease.

The basic scheme is to predict motion from picture to picture and to use Discrete Cosine Transform (DCT) to organize the redundancy. Each picture is divided into an array of macroblocks, each 16x16 pixels in size and comprising four 8x8 blocks of luminance (Y), and two 16x16 blocks of color information (U and V). The Y, U, and V information in each macroblock is then compressed using DCT encoding and motion compensation. Simply put, given the 16x16 block in the current picture that you're trying to code, you look for a close match to that block in a previous or future picture depending on the picture type. Then, with a little bit of math, you recreate the information dropped during encoding.

The elementary stream, which is the output of an MPEG encoder, contains all the necessary information to produce a decoded video. The elementary stream is composed of multiple GOPs. The pictures that make up each GOP are comprised of multiple slices, which are a group of macroblocks — each comprised of the smallest building block of the stream (the DCT coefficient blocks). The individual building blocks of the elementary stream (macroblocks, slice, picture, and GOP) all have a header, which contains all the part-specific information

Putting the pieces together

That's great, but now all we have is a silent movie! What about the audio and other content information commonly associated with streaming media? What is important to remember with the MPEG set of standards is that they're all different sets of tools that developers can use to create an entire multimedia package. Each individual standard, such as MPEG-4, has multiple parts that cover different aspects of the standard. For example, MPEG-4's parts are as follows:

Part 1: Systems
Part 2: Visual
Part 3: Audio
Part 4: Conformance Testing
Part 5: Reference Software
Part 6: Delivery Multimedia Integration Framework
Part 7: Optimized Software for MPEG-4 Tools
Part 8: MPEG-4 on IP Framework
Part 9: Reference Hardware Description
Part 10: Advanced Video Coding (AVC)
Part 11: Scene Description
Part 12: ISO Media File Format
Part 13: IPMP Extensions
Part 14: MP4 File Format
Part 15: AVC File Format
Part 16: AFX (Animation Framework Extensions) and MuW (Multi-user Worlds)

Although each of these parts is interrelated, individual parts can stand alone or be combined with any number of other parts to construct a more complex stream. When more than one type of stream, such as video, audio, and data, are combined, they create a transport stream. The systems section (part 1) of the MPEG standard provides a means for which to combine these individual streams, as well as information regarding navigation, access control, error protection, and methods for transmitting timing information to the decoder.

Applications

In the business world, videoconferencing has long been a staple among large corporations that could afford the expense of the hardware and associated "air time" call costs. With the introduction of MPEG-4 part 10/ITU H.264 AVC, the bandwidth costs associated with streaming video have been cut in half while still maintaining the same video quality and clarity. The decrease in bandwidth not only saves transmission costs for videoconferencing, but also reduces the strain on the local corporate network. Video-on-demand utilization by the corporate world has grown, particularly in the form of training videos. From their desktops, employees can launch a stream showing them how to properly file a TPS report or, more importantly, how to use the coffee maker.

The use of streaming media for advertising has also taken off in the retail market with dynamic digital signage, point-of-sale advertisements, and interactive product previews. During a recent visit to the mall, it was incredible to see all of the different implementations of media streaming — from interactive advertisements to the digital store directory. One of the more interesting setups was in a music store where a shopper could pick up any CD in the store, walk up to a listening station, scan the barcode, and then have access to a reasonable length audio clip of every track on the album. The same type of preview station was also available at a computer software store where upon scanning select video games, a preview video was shown on an adjacent screen. These systems rely on a central media server with the barcode scanners acting as the mouse for a video-on-demand interface.

Video-on-demand is one of the biggest buzzwords when it comes to streaming media because it's becoming so commonplace. With most digital cable providers offering not only video-on-demand rentals with all the shuttle controls of a VCR, but also previous episodes of popular TV series, there may be a shift in the marketplace away from the traditional analog methods for media distribution.

Satellite radio is another example of a digital media stream mimicking a traditional distribution method. XM Satellite Radio uses a combination of the MPEG-based AAC and Coding Technologies' Spectral Band Replication (SBR™) technology called CT-aacPlus, while Sirius Satellite Radio uses the PAC v4 Audio Codec, developed by iBiquity Digital. These fully digital streams are beamed down to the receivers where they're decoded. These streams also contain the associated data such as artist, song name, album, and the current score for sports streams. Traditional terrestrial radio stations are also using streaming media for management of their music libraries, allowing a DJ (which is kind of ironic to use that term in this situation) to look up songs based on differing criteria, cue it up, and play it just as he or she would a traditional format.

Streaming media for the general public exists mostly through the Internet for PC-based applications; however, that's beginning to change with the introduction of new hardware. The open structure of the MPEG standard allows manufacturers to develop new methods for increasing quality at a lower bandwidth while still remaining compatible with existing systems. These increases in the codec's efficiency will launch the next big step in streaming media, which will be cost-effective implementation in the home market. This has already begun with the introduction of digital cable set-top boxes. It won't be long before we see an affordable home media center capable of storing everything from family home movies to the collector's edition of Caddyshack and streaming them via a home network to displays with integrated decoders, replacing the defacto standards of VHS and DVD.

David A. McNell, CTS

David McNell is an AV engineer working in the Special Technologies Group of Newcomb & Boyd, Atlanta. He is a Certified Technology Specialist (CTS) and an EIT.