White Paper

Choosing the Right Web Video Codec

by Dominic Milano

By leveraging the power of Internet-based video delivery, you can extend the reach and increase the ROI of your media assets. Whether you are using the Internet to distribute interactive, video-based brand extensions of your hit television series or sharing critical business information via your corporate intranet, it is vital that your content make a good impression. For video content delivered over the web, good impressions are the result of perceived quality. And while compelling content can transcend technical imperfections, there's a lot to be said for content that looks and sounds great.

Producing professional quality web video starts with professional quality source material.

A host of factors contribute to the perception of quality, including lighting, camera angles and motion, audio levels and clarity, color consistency and so on. Viewer expectations, too, depend on the nature of the content. Consumers, for example, expect exceptional quality when watching feature film trailers or episodes of their favorite television shows online, whereas viewers are more likely to be forgiving of breaking news footage shot with a low-resolution, video-enabled cell phone. In other words, when describing web video, "quality" is a relative term defined by context.

H.264 and VC-1 are currently the most popular web video codecs. Apple QuickTime Player and Adobe Flash Player technologies support H.264, while Microsoft Windows Media Player technologies support VC-1. H.264 and VC-1 can scale to accommodate delivery across the entire bandwidth spectrum, and both codecs can deliver high-definition content at full-screen resolutions when combined with a broadband connection, a modern CPU and hardware acceleration.

But what if your target audience doesn't have access to broadband connections and the latest CPUs? Flash, QuickTime and Windows Media provide a number of codec alternatives to choose from. This document provides a high-level overview on how to select an appropriate codec based on your content, target platform and business considerations.

Table of Contents

Codec Terminology

Codec — A combination of COmpression and DECompression — words that describe what codecs do. They compress media on input and decompress it on playback. Simply put, a "web codec" is a codec that can be used to deliver media online.

Container — A technology that can house many different codecs. Popular web codec containers include Adobe Flash, Apple QuickTime and Microsoft Window Media. These containers are capable of accommodating a wide variety of media types. QuickTime, for example, can handle animation, music, speech, text, video and much more. Containers can also be called "codec containers," "media containers" or "wrappers."

Format — Comprises one or more containers and a set of codecs.

Platform — The device your content will be delivered to and viewed on.

Formats can be either platform-specific or platform-independent. Windows Media Video (WMV) is one such platform-specific format. WMV content is intended for playback on Windows Media Player, Silverlight and compatible applications.

MPEG-1 is an example of a platform-independent format. Flash is a cross-platform, platform-specific format. That is, Flash Player is compatible with a variety of operating systems, but Flash content is intended for playback on the Flash family of players, which includes Flash Player, Flash Lite for mobile playback and Adobe Media Player.

Note: *If you would like to learn about more of the technical jargon associated with media compression, the "Transcoding 101" white paper can help you quickly get up to speed. Find it at: (http://www.rhozet.com/transcoding_101.pdf)*

It's All About Delivery

Irrespective of the formats and codecs used during capture and editing, the choice of delivery codec plays a crucial role in the viewing experience.

Even the most pristine source material can be ravaged by compression choices that don't pair well with aspects of the source material.

Footage of action sports, for example, will benefit from a codec that can handle high-motion content. Medical imaging used in collaborative distance surgery calls for a codec that can run in real-time and reproduce ultra-fine details without introducing artifacts.

Other types of content may be less demanding. Tutorial footage that explains how to use a particular piece of software, for example, might be perfectly acceptable if encoded using a screen capture codec optimized for low-bitrate delivery. As the popularity of MP3 audio files has demonstrated, there is plenty of room in the market for "good enough" compression.

In some cases, business concerns might be as much or more of a priority as excellent image and audio quality. Security, for example, could be of paramount importance to a corporate officer delivering a briefing to the financial community or to a film production company posting dailies online for studio executives and other stakeholders to review and approve. In these examples, your codec choices will likely be driven by your choice of playback platform.

Prevailing bandwidth conditions also have a profound effect on the viewing experience, so codec choices must accommodate available bandwidth. High-definition, full-screen web video, for example, requires that an audience have broadband Internet connections to be effective. Broadband connections, however, are not yet ubiquitous, so using a codec that is scalable and highly efficient can help mitigate issues that arise from disparate connection speeds and playback capabilities.

Another factor that must be taken into account when choosing a web video codec is how efficiently it can be decoded. Popular codecs such as H.264 and VC-1 offer highly sophisticated filtering, motion compensation and other powerful compression features that stress all but the most up-to-date computers. It is important, therefore, to consider your target audiences' computing capabilities. This task is easier for enterprises that equip their personnel with a standardized computing platform. Anyone trying to reach a mass-market audience over the web must take a least-common-denominator approach or deliver versions of their content scaled for low-, medium- and high-bitrate playback.

These factors also come into play when choosing an appropriate transmission mode, which typically means either progressive download or real-time streaming. Which mode you choose will have ramifications across your compression workflow from encoding through delivery.

Fortunately, the dominant web video formats — ***Adobe Flash, Apple QuickTime, Microsoft Windows Media*** — support ***H.264*** and ***VC-1***, two codecs that are virtually indistinguishable from each other at high bitrates. The differentiating factors are found in subtle but important details — post-processing filtering features, the installed base of playback platforms (media players and browser plug-ins as well as compatible operating systems), security features, digital rights management capabilities and so on.

Delivering individual video files in whatever web video codec and format you choose is often simply a matter of outputting your content in that format from your nonlinear editing system. Encoding on a larger scale requires scalable, high-performance universal media transcoding technology such as Rhozet™ Carbon Coder and Carbon Server.

Web Video History

The shape of web video has always been dictated by prevailing bandwidth conditions and available CPU horsepower. In the early days of web video (circa 1995), most consumers accessed the Internet over limited-bandwidth dial-up connections using 28kbps or 56kbps modems. High-speed access was typically only available in corporate environments where T1 or T3 lines were shared among hundreds of employees. Computers were a lot less capable than they are today.

As a consequence, web video of the mid-1990s was often delivered as a downloadable file. A downloadable file could only be played after it was completely downloaded to a user's computer. QuickTime and MPEG-1 were popular codec and format choices, because they could deliver acceptable quality (for that time) at low bitrates and low resolutions. Such video played back at screen sizes that were essentially on par with that of a modern video iPod — 320x240 pixels — or smaller. Fortunately, modern codecs are much more efficient than older codecs, so they deliver much higher quality than older codecs at comparable data rates.

While broadband access has become pervasive enough for most North Americans and Europeans to take it for granted, it is important to remember that in some developing countries — South Africa, for example — dial-up modems are still prevalent.

Your codec and format choices must make different assumptions about prevailing bandwidth conditions if you are targeting audiences in disparate parts of the world.

Transmission Modes

The two most popular forms of transmitting video over the web today are ***real-time streaming*** and ***progressive download***. Each has its strengths and weaknesses. Flash, QuickTime and Windows Media are all capable of both types of transmission.

Real-time Streaming

Streaming video plays back in real time, regardless of the duration of the content.

Real-time streaming is often employed for transmitting content that is longer than about 10 minutes. To accomplish this, specialized streaming server software is required. Windows Media streaming requires server software that supports the MMS (Microsoft Media Services) protocol. QuickTime utilizes RTSP (Real Time Streaming Protocol), while Flash uses RTMPE, an enhanced version of Adobe's Real Time Messaging Protocol (RTMP). Many corporate firewalls block streaming protocols, so the major streaming solutions are able to fall back to TCP or HTTP transmission. Streaming is generally considered more secure than progressive download, however each of the major formats and their respective media players boast some form of content protection for both streaming and progressive download files.

Progressive Download

Progressive downloads are files that can start playing while they have only been partially downloaded and buffered in memory to the user's computer.

Progressive downloads can be transmitted using standard web protocols such as HTTP and FTP that don't require special servers to delivery. Progressive transmission lets you control video and audio quality at any connection speed. The tradeoff is that playback does not occur in real time.

Most progressive download files can be set to automatically start playing after a certain amount of the content has been received. In scenarios where the duration of the content is short enough, the file will finish loading before playback catches up to the download. This gives the user an experience that mimics streaming. It's more common for progressive files to be of sufficient length that playback catches up to what's been preloaded in the buffer, causing viewers to wait for the buffer to fill before playback continues. In practice, progressive downloads for files that are longer than about 10 minutes can be problematic from a viewer-experience point of view.

Similarly, so-called "Instant-On" or "Watch Instantly" streaming allows video to begin playing by reducing buffering/wait time. Instant-On experiences require broadband connections. File sizes and therefore appropriate data rates for the available bandwidth are crucial, as is the choice of codec. All three of the major formats, Flash, QuickTime and Windows Media, support this function using their respective server and player technologies.

Non-real-time playback has another benefit — two-pass variable bit rate (VBR) encoding can be used to improve video and audio quality. Most codecs function in one-pass mode, compressing the file as it is being read. Some one-pass codecs run at a constant bit rate (CBR). Other codecs offer one-pass VBR, where the data rate can vary over time, resulting in better average quality. Two-pass VBR codecs use one pass to measure the relative complexity of the content. On the second pass, the codec applies the available bits over the entire file to achieve the highest average quality, spending more bits on complex sections and fewer bits on less demanding sections. This technique produces substantial gains in compression efficiency over CBR and one-pass VBR methods.

Multicasting

Multicasting is a variation on real-time streaming in which multiple viewers can watch the same content at the same time. Multicasting is used for capturing, compressing and storing as well as webcasting live events in real time. Multicasts require that every router between the server and each client be multicast-enabled. Again, all three of the major formats support multicasting, as do most modern routers, but the function isn't always turned on. The benefit of multicasting is that a multicast stream preserves overall network bandwidth and can be useful for transmitting over low-bandwidth LANs. Intracompany communication and distance learning are two common uses for multicasting.

How a format handles quality of service is an important consideration when choosing a web video format.

Quality-of-service issues can arise with both streaming and progressive download content. Web video is delivered in packets. What happens when a packet is dropped or corrupted during transmission is an area each of the major formats handles differently. Viewers expect a seamless, trouble-free viewing experience when watching certain types of content — especially content they paid for. Any interruption of service could dissuade customers from remaining customers.

Protective Measures

Technology to restrict unauthorized use or distribution of content is an essential ingredient for many businesses. Most computer-based content protection systems use a combination of a license server and public key encryption, which ties specific content to a specific machine or device. These encryption methods reduce but do not eliminate piracy, because all broadly deployed technologies have been beaten. The digital rights management technologies used in the major web video formats are:

FairPlay (Apple, iTunes, iPhone and iPod specific). FairPlay-protected files are regular MPEG-4 container files with an encrypted AAC audio stream. FairPlay does not prevent a file from being copied, but rather controls the conditions under which it can be decoded and played back. FairPlay is built into the QuickTime architecture.

Flash DRM (Adobe, Flash specific). The Flash DRM supports encrypted streaming and protected downloads, and enables content licensing as well as limited usage or download-to-own delivery models.

Windows Media DRM (Microsoft, Windows specific). The WM DRM supports direct and indirect licensing, subscription services, pay-to-download and rental services, as well as VOD and pay-per-view models.

Business Considerations

Business considerations begin with a fundamental question: "Can your audience watch your web video?" Streaming high-definition content over a corporate intranet might seem like an ideal way to use great looking video to deliver a message to everyone in your company, but there are myriad factors that could torpedo the experience, including:

- Insufficient network bandwidth
- Content data rate and user connection speed mismatch
- Underpowered client-side computers that cannot decode the content data stream
- Users don't have a compatible browser plug-in or media player installed.

In context of the Internet at large, these factors still apply and are compounded by things over which your IT department has no control. An in-depth discussion of the possibilities is beyond the scope of this document, but it is essential to understand the typical computing power, available connection speed and operating system(s) your target audience is likely to be using. Your task may be simpler if you are only trying to reach a crew of early adopters running the latest computers on a single operating system, although using an enterprise-level transcoding system such as Rhozet Carbon Coder makes short work of delivering content in each of the major formats as well as in various legacy formats at a variety of resolutions and data rates.

Brand, Control, Track and Monetize

Flash, QuickTime and Windows Media can be played back in their respective media players or through browser plug-ins. The appearance of each of the available media players can be customized, allowing companies to deliver branded experiences to their customers. The Apple QuickTime Player and Microsoft Windows Media Player are available for both Macintosh and Windows operating systems. Both can be given custom skins. Adobe and Microsoft offer next-generation playback technologies — Adobe Media Player and Microsoft Silverlight, respectively — that are runtime-environments on which custom applications can be written to deliver a wide variety of interactive experiences.

The ability to maintain strict control over access to content, track viewer behavior, capture vital statics and monetize content in various ways are must-have capabilities for some enterprises. These capabilities are another area where there are significant differences between what each of the major web video formats brings to the table.

Adobe Media Player combined with Flash Media Server 3, for example, enables new revenue generation models including advertising that can be refreshed dynamically even after the content has been downloaded to a

Standard (In)compatibility

H.264 and VC-1 are both industry standards. However, just because a technology supports one of these "standards," it doesn't guarantee compatibility, because H.264 and VC-1 consist of many different "Profiles" and "Levels."

A **Profile** defines a subset of features needed to support a specific delivery platform. A **Level** specifies constraints within a Profile, such as maximum resolution or bitrate.

H.264 Profiles appropriate for web video applications are: Baseline Profile (BP) for limited computing power and Main Profile (MP) for typical consumer computers.

For VC-1, your choices are Low and Main Profiles at low, medium and high levels depending on target platform.

user's computer, subscription-based delivery that requires authentication on a per-user or group basis, and more.

Microsoft Silverlight supports Microsoft PlayReady DRM content protection and offers developers a programming environment for delivering rich media experiences including streaming video-on-demand, protected downloads and much more. Adobe Media Player and Microsoft Silverlight are cross-platform applications that run on Windows and Mac OS.

Formats and Supported Codecs

The following is a look at each of the major formats and the codecs they support as well as a breakdown of the types of content for which they are best suited. While some of the formats support numerous codec options, H.264 has become the video codec of choice for both QuickTime and Flash. VC-1 is Microsoft's primary video codec, but the company has announced support for H.264 in an upcoming version of Silverlight.

For audio associated with H.264 video files, AAC (Advanced Audio Coding) is the predominant audio codec, although Flash and QuickTime implement different flavors of AAC. QuickTime, for example, implements AAC-LC (Advanced Audio Coding Low Complexity). The Flash implementation of AAC supports HE-ACC v2, which is also known as HE-AAC+. HE-AAC v2 extends AAC by applying spectral band replication (SBR) and parametric stereo (PS) techniques to increase coding efficiency at low bitrates. Microsoft's Silverlight implementation of AAC will initially be for AAC-LC.

Flash (Adobe Systems, Inc.)

In its original incarnation, Flash was a small-footprint compression format designed for delivering animation over low-bitrate dial-up connections. Flash Player 6 introduced video streaming using Macromedia Flash Communications Server MX. Sorenson Spark was the codec. Spark was a moderately efficient codec based on H.263. Progressive download was added with Flash Player 7, and support for the On2 VP6 codec was added in Flash Player 8. As of Flash Player 9, the primary video codec is AVC/H.264, aka MPEG-4 Part 10. Audio support includes the audio codecs defined by MPEG-4 Part 3 — AAC Main, AAC-LC and HE-AAC v2.

Adobe offers two media player technologies: Flash Player and Adobe Media Player. Flash Player is a browser extension which, according to a study conducted by Millward Brown[1] for Adobe, is installed on 99% of all Internet-connected computers. Flash Player runs on a variety of operating systems including Windows, Mac OS and Linux. Flash Player supports playback of FLV (FLash Video) and SWF (ShockWave Flash) files as well as MPEG-4 (MP4/F4V) content when a user is online. Adobe Media Player is a desktop application that runs on Windows and Mac OS and supports playback of streamed or downloaded FLV and MPEG-4 content when users are online or offline.

1. http://www.adobe.com/products/player_census/flashplayer/

Flash Video Codecs

H.264, aka AVC (Advance Video Coding) or MPEG-4 Part 10, is a highly efficient codec designed to produce better quality at substantially lower bitrates than older codecs such as H.263 and MPEG-1. Flash supports the following profiles: Baseline, Main Profile (MP), High Profile (HiP), and High 10 Profile (Hi10P). Baseline applications include videoconferencing and mobile running on limited computing power. Main Profile and High Profile are primarily for broadcast applications. High 10 Profile adds 10-bit color subsampling to HiP. Supported levels/resolutions range from 160x112 standard definition video to 1920x1080 high definition video. Typical frame rates range from 5 to 30 fps. Higher resolutions and frame rates require more computing horsepower to decode. Flash 9 supports hardware acceleration for full-screen playback. The H.264 codec is appropriate for compressing all types of video content.

On2 VP6 was the primary video codec in Flash 8. It is still supported in Flash Player 9. Supported levels/resolutions are the same as those for H.264 in Flash. VP6 supports both one-pass CBR and two-pass VBR encoding. Use VP6 if your target audience is using computers that aren't capable of handling the H.264 codec either because the CPUs are not up to the task, your audience doesn't have Flash Player 9 or later installed or they are running older operating systems.

Flash Audio Codecs

Flash Player supports the AAC Main, AAC-LC and AAC-HE v2 audio codecs defined by MPEG-4 Part 3. AAC stands for Advanced Audio Coding; LC indicates Low Complexity and HE indicates High Efficiency. AAC Main is based on the MPEG-2 version of AAC and adds perceptual noise shaping to improve quality at lower bitrates. AAC Main supports up to 5.1-channel audio. AAC-LC is less efficient than AAC Main, but requires less computing power to encode and decode. AAC-LC is optimized for streaming applications such as Internet radio. HE-AAC v2 enhances frequency domain coding efficiency and improves low-bitrate stereo signals. HE-AAC v2 can handle up to 48 audio channels and 5.1 as well as 7.1 surround sound.

MP3, aka MPEG-1 Layer 3, was the primary audio codec of Flash Player 6. It is optimized for 160kbps data rates at 8kHz.

Speech is a licensed version of the Nelly Moser codec that is best suited for low-bitrate, low-fidelity spoken-word, monaural content.

ADPCM and RAW offer good quality encoding with little or no compression, respectively. Neither is particularly well suited for web delivery.

H.264 vs. VC-1

At high bitrates, H.264 and VC-1 are essentially indistinguishable:

- Under high bandwidth stress, H.264 tends to degrade to softness instead of blockiness, whereas VC-1 tends to retain more detail, but can show more artifacts due to its stronger in-loop deblocking filter. At web/download bitrates, both codecs offer great quality.
- VC-1 has an optional post-processing filter that can be used to reduce artifacts when sufficient CPU horsepower is available. This filter is implemented in Windows Media Player.
- H.264 tends to take more CPU power to decode at maximum complexity. Comparing VC-1 Advanced Profile and H.264 High Profile, each with all of its options turned on, VC-1 can decode about twice as many pixels per MIPS. This has more bearing on HD-resolution content than lower resolution content and doesn't matter at all when there's a hardware decoder. When constrained by software decoder performance, VC-1 can handle Advanced Profile, while H.264 is limited to Baseline, where VC-1 outperforms H.264 in quality/efficiency. This has become an important consideration in recent years, because consumer broadband access has outpaced the hardware upgrade cycle of many households and businesses.

QuickTime (Apple Computer, Inc.)

QuickTime was introduced in 1991. It is known for supporting an expansive set of codecs suitable for nearly every media type. In its current incarnation (QuickTime 7), H.264 aka MPEG-4 Part 10 or AVC (Advanced Video Coding) is its web video codec of choice; AAC-LC (Advanced Audio Coding Low Complexity) is its corresponding audio codec. These are the video and audio compression formats of the popular iPod and the iTunes Music Store.

The QuickTime Player is preinstalled on all Macintosh computers and a version is available for Windows machines. Compatibility between the two versions is essentially at parity. QuickTime Player 7 supports playback of up to full-screen, high-definition content.

QuickTime support for progressive download predates its support of RTSP streaming, which was introduced with QuickTime 4 in 1999. QuickTime streaming requires the Apple QuickTime Streaming Server (QTSS) or the open-source Darwin Streaming Server. QTSS supports streaming live and on-demand content.

QuickTime Web Codec Support

H.264, aka AVC (Advance Video Coding) or MPEG-4 Part 10, is a highly efficient codec designed to produce better quality at substantially lower bit rates than older codecs such as H.263 and MPEG-1. QuickTime implements H.264 BP (Baseline Profile for low-latency devices such as cell phones) and HiP (High Profile) for high-definition broadcast and Blue-Ray Disc applications. The hard-to-decode features of H.264 (in particular, CABAC entropy coding and multiple reference frames) are turned off in the QuickTime implementation of the codec. This sacrifices some efficiency to achieve faster decoding performance and facilitates instant-on streaming.

CABAC, in case you're wondering, stands for context-adaptive binary arithmetic coding. CABAC is a lossless compression technique that's highly efficient. CABAC takes more power to decode than its lower complexity alternative, CAVLC (context-adaptive variable-length coding), which is still more efficient than techniques used in older codecs.

H.264 is appropriate for compressing just about any type of video content. One of the codec's strengths is scalability, so you can use it to deliver video at a variety of resolutions and bitrates to a variety of devices. Lower bitrate H.264 (~40kbps to 300kbps) can be played on modern cell phones and consumer-level computers. Delivery of H.264-encoded files to other mobile devices such as an iPod Touch can use resolutions of up to 640x480 at data rates of up to 1.5mbps. Higher-definition H.264 files require more computing muscle to decode, so your audience will need to be on sufficiently modern computers and broadband connections to view such content. For an up-to-date list of QuickTime 7 Macintosh and Windows system requirements for H.264 playback at a variety of resolutions and frame rates, see http://www.apple.com/quicktime/guide/hd/recommendations.html

Note that QuickTime streaming requires that H.264-encoded files be hinted. Hinting gives the streaming server information it needs to properly stream the media. Prebuilt playlists of AAC-LC-encoded audio files, however, don't require hinting. Rhozet Carbon Coder takes care of hinting automatically during the transcoding process.

MPEG-4 Part 2 is the original MPEG-4 video codec that's been largely superseded by AVC/H.264, which provides up to four times the frame sizes at any given data rate. MPEG-4 video codec support in QuickTime 6 was for the MPEG-4 Part 2 codec. This codec is still supported in QuickTime, but is more useful for video delivery to mobile phones than video on the web.

H.263 is a low-bitrate codec originally designed for teleconferencing. If your audience is using older computers that are not able to decode H.264, H.263 is a viable option. H.263 is limited to native resolutions of 352x288, 176x144 or 128x96. Encoding at other resolutions causes the codec to select the closest supported resolution and scale up on playback. H.264, however, provides much better quality at comparable data rates and is well suited for video conferencing applications. Apple iChat, for example, relies on the H.264 codec.

H.262, aka MPEG-2, has been the codec of choice for standard-definition DVD and broadcast television. It supports interlaced video, a feature that is unnecessary for web-based delivery of video to computer screens. H.262 file sizes are impractical for web delivery of any sort.

MPEG-1 predates the MPEG-2 standard and is still in use today. MPEG-1 video was originally intended for compressing digital video and CD audio down to 1.5mbps at 26:1 and 6:1 compression ratios, respectively, for delivery on VideoCD. MPEG-1 does not support interlaced video, making it a logical choice for web video applications. Two-pass VBR encoding can help mitigate its relative lack of compression efficiency. For a variety of reasons not the least of which is that there are no licensing fees required for implementing MPEG-1 encoding or decoding, MPEG-1 video and audio (including MP3, aka MPEG-1 Layer 3 audio) are among the most widely supported compression formats in the world. If your target audience uses Unix-based computers, progressive download MPEG-1 files offer a viable compression option.

Sorenson Video v3 (SV3) was at one point the default QuickTime video codec. It has been built into QuickTime since v5.02. The SV3 codec works well for progressive download content, and using two-pass VBR can help achieve optimal quality. When streamed with the QTSS, SV3 can dynamically drop B-frames if there's not enough bandwidth to support a given data rate. (B-frames are bidirectional prediction frames used to improve compression efficiency.)

QuickTime Audio Codecs

AAC-LC is the audio codec of choice for the most recent incarnation of QuickTime (v7.5.5 as of this writing). Advanced Audio Coding Low Complexity is one of the audio codecs in the MPEG-4 suite of standards. AAC-LC provides much better compression efficiency than MP3 (MPEG-1 Layer 3) and includes a native packetizer for RTSP streaming. The codec is appropriate for speech and music. Complex program material benefits from higher sampling and data rates.

QuickTime 7 supports multichannel audio in MOV files, such as those used in high-definition movie trailers on Apple.com/trailers/. The Apple implementation of the MP4 container, however, supports only mono and stereo audio, so multichannel MOV files need to be recompressed when output as MP4 files. Rhozet Carbon Coder enterprise transcoding software handles this transcoding task automatically.

MP3, aka MPEG-1 Layer 3 audio, is still supported in QuickTime. Its implementation does not include a native packetizer, so MP3 is best used for progressive download playback. MP3 provides acceptable quality for both music and speech, and served as the original proof-of-concept for "good enough" compression in the hay day of music file sharing services.

Windows Media (Microsoft Corporation)

Windows is most often associated with Intel-based computers running Microsoft's Windows operation system. Windows Media Player is included in the Windows OS and is available for Mac OS. It handles streaming and progressive download content, as well as live broadcast and on-demand video. Windows Media offered integrated digital rights management (DRM) features before either QuickTime or Flash. Its DRM technology has been evolving steadily since it was introduced in 1999. All Windows Media Series codecs support the Windows Media Digital Rights Management platform to distribute securely packaged copy-protected digital media.

The Windows Media Player and Silverlight support a limited number of codecs. Principal among them is VC-1 (aka SMPTE 421M), which is a highly efficient codec capable of delivering high-definition, interlaced or progressive-scan video on any medium to any compatible device. Microsoft has announced support for H.264 in an upcoming release of Silverlight. The Windows Media Video (WMV) and Windows Media Audio (WMA) codecs are backwardly compatible with earlier versions of the Windows Media Player and compatible technologies.

Windows Media Video Codecs

Windows Media Video 9 (WMV9) is Microsoft's implementation of the VC-1 codec. It supports Simple, Main and Advanced profiles. Simple and Main support a wide range of bitrates appropriate for high-definition content to web video delivered at low-bitrate, dial-up connection speeds. WMV9 supports two-pass VBR encoding. Advanced profile VC-1 is transport-independent, meaning it can be used to deliver content to non-Windows Media-based systems including wireless networks using real-time transfer protocol (RTP). Microsoft claims Advanced profile achieves comparable quality at data rates as low as one-third that of the MPEG-2 codec. Like H.264, VC-1 can be used to compress all manner of streaming and progressive download video content, from podcasts to high-definition movies on demand. WMV9 is compatible with Windows Media Player 6.4 or later; Advanced profile playback requires Windows Media Player 7 or later.

Windows Media Video 9 Image Version 2 is used to turn still images into video using various effects such as pan, zoom and crossfades between clips. Data rates as low as 20kbps can be used to deliver the files. CBR and one-pass VBR encoding is supported. WMV9 Image Version 2 is compatible with Windows Media Player 7 or later. Content encoded with WMV9 Image Version 2 is not compatible with its earlier incarnation, the WMV9 Image codec.

Windows Media Video 9 Screen is optimized for compressing sequential screenshots and relatively static video captured from computer displays. The codec leverages the lack of motion to achieve good quality at very high compression ratios. The codec is useful for encoding tutorial and demo videos and delivering content to modestly powered CPUs. WMV9 Screen is compatible with Windows Media Player 7 or later.

Windows Media Legacy Codecs

Windows Media Video 8 provided competitive quality at a full range of data rates and scaled well, providing good playback on under-powered CPUs as well as the fastest computers of its day. It offered one-pass CBR, two-pass CBR, one-pass VBR and two-pass VBR encoding. Two-pass CBR used an analysis pass to achieve higher average quality at a constant bitrate, a mode that was well suited for real-time streaming.

Windows Media Video 7 is less efficient than WMV 8, delivering less quality at a given bitrate. WMV 7, however, is faster to encode than WMV 8. Use this codec to deliver streaming or progressive download video if your audience uses older computers.

Windows Media Audio Codecs

Windows Media Audio 9 (WMA9) is a codec that is backwardly compatible with earlier versions of the Windows Media Player, so it can be used even if your target audience is running Windows Media Player 6.4 or later. The codec uses 16-bit audio and samples at 44.1kHz and 48kHz. Microsoft claims WMA9 can produce CD-quality audio at data rates from 64kbps to 192kbps and yields audio that's 20 percent better than audio sampled with the Windows Media Audio 8 codec at equivalent data rates. VBR encoding is supported.

Lossy Codec Basics

Web video codecs are "lossy" codecs, that is, they discard redundant information to reduce file size in a way that produces acceptable image quality. To do that, they utilize a technique that relies on the fact that most video frames are similar to the frames that surround them. By comparing the current frame to one or more earlier or later frames, an interframe codec can treat redundant image information as it would a proxy file. That is, instead of reproducing the redundant information bit-by-bit, the codec can simply point to or reuse the previous frame's redundant content and send only those portions of the content that have changed within the new frame or between frames. This technique is most effective with fairly static content such as talking heads.

To handle high-motion content, codecs such as H.264 and VC-1 use motion estimation and compensation techniques to increase compression efficiency and maintain image quality. When these codecs look at a frame, they use a technique called block motion compensation (BMC). That is, they examine square-shaped groups of pixels called a "macroblock" within a reference or "I" frame.

Older codecs such as H.261 and MPEG-1 video required fixed-sized macroblocks. H.264/AVC and VC-1 use variable blocksize motion compensation (VBSMC), which allows the encoder to dynamically select the size of the blocks. Using larger block sizes can reduce the number of bits needed to represent motion vectors.

The downside of BMC is that it can result in blocking artifacts (sharp edges) and ringing (closely spaced ghosts or alternating dark and light areas along vertical or diagonal edges).

Windows Media Audio 10 Professional (WMA 10 Pro) supports a full spectrum of possible resolutions for low-bitrate mobile delivery to full-resolution 24-bit/96kHz stereo, 5.1 and 7.1 surround sound. It is flexible enough to be used to compress audio for playback on computers with high-fidelity audio hardware and for delivering audio on mobile devices. WMA 10 Pro can do streaming, progressive download or download-and-play delivery at 128kbps to 768kbps. Files encoded for multichannel, high-resolution (24-bit/96kHz) playback that are played back on a system that cannot handle multichannel surround sound will playback as 16-bit stereo audio.

Windows Media Audio 9 Lossless, as its name implies, is a lossless codec that creates a bit-for-bit duplicate of the original file, which makes it most useful as an archival format. Content is typically compressed at a 2:1 or 3:1 ratio depending on its complexity.

Windows Media Audio 9 Voice is a low-bitrate codec used for compressing speech. It can also be used to compress content that includes both voice and music. In speech mode, the codec leverages the narrow frequency range and relative simplicity of the human voice to achieve optimal results. In music mode, the codec behaves like the WMA9 codec. Encoded content can be configured to automatically switch between voice and music modes. WMA9 Voice is well suited for streaming of Internet radio, podcasts, e-books and voiceovers at bitrates as low as 4kbps at 8kHz.

MP3 and the Windows Media Audio codecs listed above are supported in Silverlight. AAC-LC support has been announced for Silverlight.

Legacy Formats and Codecs

RealMedia (RealNetworks, Inc.)

RealNetworks was the first company to offer real-time streaming technology. When RealPlayer was introduced in 1995, it was primarily known as a streaming audio application. Streaming video support was added in 1997. The RealMedia format is often referred to as "RealVideo," even if the content consisted of nothing but audio. RealNetworks SureStream technology required the RealNetworks Helix server and created multiple audio and video streams tailored to accommodate various network conditions including available bandwidth and packet-loss rate. Early versions of the RealVideo codec were based on H.263. RealVideo 8, 9 and 10 used a proprietary video codec developed by RealNetworks. RealVideo 10 (c. 2003) enabled delivery of rich media at any bitrate on any device.

RealVideo Codecs

RealVideo 10 (RV10) is 30 percent more efficient than RealVideo 9. Real Networks claims RV10 delivers the same quality as H.264 at 15 percent lower bitrates. RV10 is compatible with RealVideo 9 decoders and runs on a variety of operating systems and CPUs, including Windows, Windows CE, Mac OS, Linux and Symbian. RV10 can handle both progressive-scan and interlaced video content. One- and two-pass CBR, VBR and Quality-Based Encoding are supported. The latter maintains a constant level of quality throughout the file being encoded, without regard for bit usage. RV10 supports playback of video on low-bitrate cell phones and mobile devices as well as high-bitrate HD 720p and 1080i resolutions.

RealVideo 9 (RV9) offers significantly better compression efficiency than RealVideo 8 (RV8), although it requires a more powerful CPU to decode. Like RV8, RV9 supports two-pass VBR encoding. Its decoder can play content at frame rates higher than they were encoded at if the CPU exceeds the minimum system requirements. This is accomplished by interpolating motion vectors to generate new frames between existing frames.

RealAudio Codecs

RealAudio 8 (RA 8 and RA8) is optimized for delivering stereo audio at low bitrates. A version of Sony's ATRAC-3 (RA 8 — note the space between "A" and "8") codec that is enhanced to allow RSTP streaming can be used for 96kbps and higher data rates. For bitrates below 96kbps, RealAudio 8 uses a proprietary codec, called RA8 (no space), developed by RealNetworks.

Voice is a speech codec optimized for data rates up to 64kbps.

Music is RealNetworks' audio codec for monaural content, used for 6kbps to 64kbps monaural music clips.

Miscellaneous

MPEG-1 predates the MPEG-2 standard and is still in use today. MPEG-1 video was originally intended for compressing digital video and CD audio down to 1.5mbps at 26:1 and 6:1 compression ratios, respectively, for delivery on VideoCD. MPEG-1 does not support interlaced video, making it a logical choice for web video applications. Two-pass VBR encoding can help mitigate its relative lack of compression efficiency. For a variety of reasons not the least of which is that there are no licensing fees required for implementing MPEG-1 encoding or decoding, MPEG-1 video and audio (all three variants including MP3, aka MPEG-1 Layer 3 audio) are among the most widely supported compression formats in the world. If your target audience uses Unix-based computers, progressive download MPEG-1 files offer a viable compression option.

DivX uses the MPEG-4 Part 2 video codec and is most associated with ripping content to DVDs.

About Rhozet

Since 2004, Rhozet, a business unit of Harmonic Inc, has focused on designing scalable, high-performance, universal media transcoding technology for delivering content in any format, at any time, on any device, smoothly, efficiently and in the most cost-effective manner possible. Rhozet began as part of Canopus and operated as an independent company from 2005 until August 2007, when Rhozet was acquired by Harmonic, a manufacturer of enterprise-class hardware encoders. In addition to its enterprise transcoding products, Carbon Coder and Carbon Server, Rhozet is the developer of the popular desktop transcoding applications ProCoder and ProCoder Express, which have been marketed and sold under the Grass Valley brand since its acquisition of Canopus.

About the Author

Dominic Milano is the principal of DM&C, a company that provides content creation and consulting services in a variety of markets, including video, music and sound design, game development, interactive design, 3D modeling and animation and related creative fields. Dominic has over 30 years of experience in print, online and event media production, working on *DV* magazine, DV.com, DV Expo, *Game Developer* magazine and the Game Developer Conference, *Keyboard* magazine, *Guitar Player* magazine and more.

Special thanks to Ben Waggoner for his assistance in putting together this document.

Summary

Ultimately, choosing an appropriate web video format and codec is a matter of assessing the nature of your content, determining your business goals and evaluating the bandwidth and computing power available to your target audience.

As of this writing, there is no single, silver bullet web video codec solution, although H.264 could soon claim that title. Both QuickTime 7 and Flash Player 9 support H.264, and soon Microsoft Silverlight will too. H.264 is highly efficient and scalable enough to handle delivery across the entire bandwidth spectrum. VC-1 shares those attributes, however, its adoption has been limited to Microsoft-centric technologies.

If your audience does not have access to modern CPUs and broadband connections, each of the major formats offers alternative codec choices for just about any type of content in all manner of bandwidth conditions. Regardless of which format(s) and codec(s) best suit your needs, employing highly scalable, universal media transcoding technology such as Rhozet Carbon Coder and Carbon Server gives you the ability to handle high-volume, complex workflows for delivering video and audio content in any web video format at any given resolution and frame rate in a cost-effective manner.

Resources

Intraframe versus interframe compression
http://en.wikipedia.org/wiki/Video_compression

Motion compensation
http://en.wikipedia.org/wiki/Motion_compensation

FairPlay DRM
http://en.wikipedia.org/wiki/FairPlay

Flash DRM
www.adobe.com/products/flashmediarightsmanagement/features/

http://www.adobe.com/devnet/flashmediaserver/articles/digital_media_protection.html

Windows Media DRM
http://www.microsoft.com/windows/windowsmedia/forpros/drm/default

mspxhttp://en.wikipedia.org/wiki/Windows_Media_DRM

H.264 overview
http://en.wikipedia.org/wiki/H.264

VC-1 overview
http://www.microsoft.com/windows/windowsmedia/howto/articles/vc1techoverview.aspx

WM and VC-1 in Silverlight 2.0
www.microsoft.com/belux/interactive/newsletter/08-05/articles/mediatechnologies.aspx