



## Designing Web Audio

By Josh Beggs & Dylan Thede

January 2001

1-56592-353-7, Order Number: 3537

398 pages, \$34.95

---

## Chapter 5 Introduction to Streaming Media

Internet streaming media changed the Web as we knew it-- changed it from a static text- and graphics-based medium into a multimedia experience populated by sound and moving pictures. Now streaming media is poised to become the de facto global media broadcasting and distribution standard, incorporating all other media, including television, radio, and film. The low cost, convenience, worldwide reach, and technical simplicity of using one global communications standard makes web broadcasting irresistible to media publishers, broadcasters, corporations, and individuals. Businesses and individuals once denied access to such powerful means of communication are now using the Web to connect with people all over the world.

The remarkable technology that allows a web site visitor to click on a button and seconds later listen to a sporting event, tradeshow keynote, or CD-quality music is the result of a rather simple but powerful technical innovation--*streaming media*. Streaming works by first compressing a digital audio file and then breaking it into small packets, which are sent, one after another, over the Internet. When the packets reach their destination (the requesting user), they are decompressed and reassembled into a form that can be played by the user's system. To maintain the illusion of seamless play, the packets are "buffered" so a number of them are downloaded to the user's machine before playback. As those buffered or preloaded packets play, more packets are being downloaded and queued up for playback. However, when the stream of packets gets too slow (due to network congestion), the client audio player has nothing to play, and you get the all-too-familiar drop-out that every user has encountered.

## Streaming protocols

The big breakthrough that enabled the streaming revolution was the adoption of a new Internet protocol called the User Datagram Protocol (UDP) and new encoding techniques that compressed audio files into extremely small packets of data. UDP made streaming media feasible by transmitting data more efficiently than previous protocols from the host server over the Internet to the client player or end listener. More recent protocols such as the RealTime Streaming Protocol (RTSP) are making the transmission of data even more efficient.

UDP and RTSP are ideal for audio broadcasting since they place a high priority on continuous streaming rather than on absolute document security. Unlike TCP and HTTP transmission, when a UDP audio packet drops out, the server keeps sending information, causing only a brief glitch instead of a huge gap of silence. TCP, on the other hand, keeps trying to resend the lost packet before sending anything further, causing greater delays and breakups in the audio broadcast.

Prior to UDP and RTSP transmission, data was sent over the Web primarily via TCP and HTTP. TCP transmission, in contrast to UDP and RTSP transmission, is designed to reliably transfer text documents, email, and HTML web pages over the Internet while enforcing maximum reliability and data integrity rather than timeliness. Since HTTP transmission is based on TCP, it is also not well-suited for transmitting multimedia presentations that rely on time-based operation or for large-scale broadcasting.

Later in the chapter, you will learn why protocols are important. Some streaming technologies such as RealAudio and Windows Media utilize dedicated servers that support superior UDP and RTSP transmission. Other formats such as Shockwave, Flash, MIDI, QuickTime, and Beatnik are primarily designed to stream from a standard HTTP web server. While these formats are cheaper and often easier to use since they do not require the installation of a new server, they are typically not used in professional broadcasting situations that require the delivery of hundreds or thousands of simultaneous streams.

HTTP streaming is thus referred to as pseudo-streaming, since technically it is possible to stream via HTTP. But it is much more likely to cause major packet drop-outs, and it cannot deliver nearly the same amount of streams as UDP and RTSP transmission. Herein lies the difference between most low-end solutions and more professional broadcasting solutions that require dedicated servers and extra bandwidth and server capacity.

## Lossy compression

Regardless of the advances in UDP and RTSP transmission protocols, streaming media

would not be possible without the rapid innovation in encoding algorithms or codecs that compress and decompress audio and video data. Uncompressed audio files are huge. One minute of playback of a CD-quality stereo audio file requires 10 MB of data, approximately enough disk space to capture a small library of books or a 200-page web site.

Standard modem speed connections--including cable modems and xDSL systems--do not have the capacity to deliver pure, uncompressed CD-quality 16-bit, 44.1 kHz audio. In order to stream across the limited bandwidth of the Web, audio has to be compressed and optimized with *codecs*, which are compression-decompression encoding algorithms. In general, compression schemes can be classified as "lossy" and "lossless."

*Lossy compression* schemes reduce file size by discarding some amount of data during the encoding process before it is sent over the Internet. Once received on the client side, the codec attempts to reconstruct the information that was lost or discarded. The benefit to this sort of compression lies in the smaller file size that results from discarding the "lost" information. The JPEG image format uses lossy compression to sample an image and discard unnecessary color information. Similarly, lossy audio compression discards frequencies on the high and low end of the spectrum and attempts to locate and remove unnecessary audio data. The technique is often referred to as "perceptual encoding" since the user is unlikely to notice the absence of this information. Lossy compression offers file savings on the order of 10:1.

Since small file size is so important on the Internet, practically all of the formats we're interested in employ lossy compression. Here's how it works. First, the client player decompresses the audio file as it downloads to your computer. Then it fills in the missing information according to the instructions set by the codec. To illustrate why lossy compression is so crucial, consider the phrase, "Now is the time for all good men to come to the aid of their country". One way to compress this would simply be to remove all the vowels and spaces: "Nwsthmfrllgdmntcmthdfthrcntry".

That cuts the message from 71 characters to 31, a 56% file savings, but of course our compressed message is unintelligible. Imagine that our codec, however, has appropriate rules for decompressing this message with minimal distortion. The conversion likely wouldn't be perfect, but it would be good enough to understand the message, something like, "Now's tha ti'm for oll gudm en to com to the aad of their country".

This is exactly what happens with lossy audio compression. The compressed file is unintelligible to the listener; the decompressed file is intelligible but of a lower quality than the original.

For example, a RealAudio speech file encoded from a standard AIFF or WAV file is

generally one-tenth the size of the original file after encoding. To reduce that file's size, first you preserve the integrity of the 1,000 Hz to 4,000 Hz frequency spectrum of the human voice and then discard the frequencies above and below those ranges. By eliminating the unnecessary low- and high-end frequencies, the encoder is able to reduce the file size while maintaining speech intelligibility. It should be noted that speech tends to have aural characteristics (sound) that extend into the 7,000 Hz range. When the area between 4,000 Hz and 7,000 Hz is reduced or removed entirely, encoded speech will sound intelligible, but it may lose clarity and sound unnatural. Furthermore, since some voices and sounds often reach into even higher frequency ranges, lossy compression and encoding can result in dull, muted, or abrasive sounds.

## Lossless compression

In contrast, *lossless compression* squeezes data into smaller packets of information without permanently discarding any of the data. Instead of permanently discarding information, lossless compression discards it temporarily but provides a "map" with which the codec can reconstruct the original file. Lossless compression results in superior audio quality, but lower compression rates.

In the lossy example, our codec had some general rules for reconstructing the message-- basically to add vowels and spaces in order to form English words. It wasn't perfect because it didn't know which English words to choose, and it wasn't always sure where one word ended and the next began.

Lossless codecs, on the other hand, are perfect. To reconstruct our message perfectly, however, would mean having a much more sophisticated set of rules. A lossless text codec would have to reproduce not only words but sensible phrases. It would have to be able to break words correctly. And it would have to have a mastery of the English language's inconsistent spelling patterns. It would in fact be, as the computer scientists say, a nontrivial endeavor.

The same goes for lossless audio codecs. They are difficult to develop (and thus expensive to license), they require substantial computing power on the user's machine, and the file savings are not as great as with lossy compression. Sadly enough, it appears that for the current time, lossy compression is necessary for knocking large audio files down to Internet-appropriate size. The good news is that lossy compression schemes are becoming more advanced, and over time the differences will become less and less noticeable to the human ear.

Now that we have discussed lossy and lossless compression and the types of protocols that enable the efficient delivery of compact audio files across the Internet, let's review the audio formats available on the market. Most of these formats will be discussed in greater

detail in the rest of the book.

## Streaming media formats

There are currently more than a dozen formats for streaming audio over the Web, from widely used formats, such as RealNetworks' RealAudio, streaming MP3, Macromedia's Flash and Director Shockwave, Microsoft's Windows Media, and Apple's QuickTime, to more recent entries that synchronize sounds with events on a web page, such as RealMedia G2 with SMIL and Beatnik's Rich Music Format (RMF). Also included are a host of downloadable formats, including Liquid Audio, MP3, MIDI, WAV, and AU.

While the high quality of MP3 has sent shockwaves through the recording industry, streaming formats like RealAudio remain the dominant audio technology on the Web right now. Indeed MP3 is being folded into multimedia streaming formats like QuickTime and Windows Media.

Throughout this book, we take an in-depth look at many of the more prevalent streaming formats. However, in this chapter, we will review all the streaming formats on the Web, including Windows Media and QuickTime, which are not featured in later chapters.

### RealMedia and RealAudio

RealMedia is the most widely adopted streaming media format on the Web. Its popularity is due in large part to the fact that it was the first streaming technology on the market. But it's popular also because of RealNetworks' laser focus on ease of use, deployment of a wide palette of developer tools, continuous support for the latest multimedia technologies, and support for both Windows and Unix platforms. RealMedia is the format of choice for professionals who want advanced controls for serving, tracking, and managing large numbers of audio streams. RealNetworks has been a trailblazer in making advanced server features, which were once accessible only to those with advanced programming skills, available to the public.

And RealMedia is likely to attract more fans as web developers begin to use the RealSystem G2 and SMIL to stream synchronized multimedia presentations over the Web. G2's major advance is the ability to simultaneously stream multiple media types as separate files instead of as one RealMedia-encoded file. This makes updating multimedia content easier, since you can simply upload one element of a presentation instead of re-encoding the whole media file.

Perhaps the most powerful feature of RealSystem G2 is RealNetworks' server architecture. Broadcasting audio with a dedicated RealServer provides the following advantages over HTTP pseudo-streaming from a standard web server:

### Bandwidth negotiation

Ensures that all users receive the appropriate encoded content for the best audio quality at their available bandwidth, from slower analog modems to faster cable or xDSL connections. RealSystem G2's new SureStream technology is even more efficient than bandwidth negotiation. SureStream can dynamically change data rates midstream to accommodate fluctuating bandwidth.

### Robust RTSP transmission

Detects and compensates for lost packets, maintaining smooth, continuous audio playback--something that HTTP streaming can't deliver.

### Splitting

Allows for splitting and routing the audio signal from one RealServer to other RealServers located at different points across the Internet.

### Clustering

Allows multiple RealServers to be clustered together so they work as a single, multiprocessor machine.

### IP multicasting

Allows all users of a network to listen to a single live stream, making efficient use of network resources. Multicasting avoids delivering numerous simultaneous point-to-point connections by broadcasting one stream to a certain point in the network where other users are requesting the same file. Multicasting is ideal for reducing server load and bandwidth congestion during live broadcasts.

While RealMedia's powerful server-side architecture supports and manages robust streaming to large audiences, this core strength results in limited interactivity. Like Windows Media and other server-side streaming technologies, RealMedia waits for a request from a listener's browser before it begins to stream media files. This helps RealMedia negotiate bandwidth congestion on the fly by sending an appropriate size stream that matches the listener's real bandwidth. But it also produces a significant time gap of a few seconds between the listener's request and the response from the server. This small time gap is inconsequential with long-playing video and audio files, but it prohibits the use of interactive sound effects such as button rollovers, sound transitions from one page to another, and loops that must respond instantaneously to a mouse click.

Thus, RealMedia is inappropriate for high impact presentations with interactive sound effects and loops. Despite significant advancements in RealSystem G2, RealAudio still trails Flash and Director Shockwave when it comes to smooth playback of high-impact interactive multimedia. High-powered interactive media requires a client-side solution

such as Flash, Shockwave, or Beatnik.

## Windows Media Technologies (Netshow)

Microsoft's Windows Media Technologies for NT/Windows 2000 includes a comprehensive suite of authoring tools and streaming services for delivering audio, video, animation, and other multimedia over the Internet. Windows Media comes with a complete set of tools for encoding and authoring streaming content including Windows Media T.A.G. Author, a utility for arranging media elements along a timeline. Windows Media presentations are played back with the Windows Media Player, which plays most local and streamed media file types including Advanced Streaming Format (Windows' native file format), MPEG, WAV, AVI, QuickTime, and RealAudio/RealVideo. Since Media Player is distributed with Windows, it has widespread distribution.

If you need a Windows NT 4.0-based solution, Windows Media Services offers several advantages:

- The Windows Media Server comes free with unlimited streams with Windows NT Server 4.0 and later.
- It allows for better playback over machines running Windows. To enable smooth multimedia playback over the Web and avoid the problematic issue of cumbersome plug-in downloads altogether, Microsoft is moving towards integrating Windows Media Player, along with Internet Explorer, directly into the Windows operating system.
- Windows Media Server integrates with Microsoft Site Server to enable pay-per-view and pay-per-minute billing capabilities, usage analysis reporting, and personalized ad insertion.
- Tools for tracking behavior are tightly integrated with the Windows NT Event Viewer and Performance Monitor, making it easy for seasoned NT administrators to manage the Windows Media Server.
- For multimedia content developers, Microsoft provides helpful authoring tools. Creating a slide show of images with synchronized audio can be accomplished by using the Windows Media T.A.G. Author.

Compared to RealMedia, however, Windows Media has some serious drawbacks:

- It runs only on Windows NT/2000. Many developers have reported problems with the stability of Windows NT for mission-critical applications such as 24-hour live

broadcasts. This can be a show-stopper for those who demand the stability of Unix or Linux servers. In contrast, RealNetworks supports NT as well as Linux, FreeBSD, Solaris, and IRIX.

- It does not support Macromedia Flash or the Synchronized Multimedia Integration Language (SMIL) standard, both of which are supported by RealNetworks.

There are also some key differences in the way Windows Media and RealMedia encode and deliver multimedia content. With RealMedia, you can create multimedia presentations by using the SMIL markup language to connect various media elements together. These media elements are encoded as separate files: RealAudio, RealVideo, RealPix, RealText, QuickTime, MPEG, and so on. The RealServer, much the same way a standard web page is served up and delivered, then streams the presentation as separate media files held together by SMIL.

"Since G2 developers are creating multimedia presentations rather than simply encoding audio or video streams, the format has a new level of complexity," says Leah Goldberg, G2 media producer for CMPnet. "However, web developers have long been familiar with the flexibility and convenience of this approach to media delivery. The challenge with G2," Goldberg claims, "is working out the timing in the component RealPix, RealText, and RealFlash files. Since the idea is to synchronize all the different media elements together, working out the sub-timing issues within each of the component files can be quite complex."

In contrast, Windows Media wraps all media elements into one Active Streaming File (ASF), Microsoft's proprietary streaming media format. According to Microsoft, with ASF any object can be placed into an ASF data stream, including audio and video, scripts, ActiveX controls, and HTML documents with T.A.G. Author. This approach, similar to Flash and Shockwave movies, provides less flexibility in terms of updating and serving content, but it offers more stable client-side playback of various media elements and tighter authoring controls. For more information about creating ASF content, visit the Microsoft web site. Microsoft provides free code to members of its Developer Network.

## QuickTime

Apple Computer's QuickTime enables the delivery and playback of video, audio, animation, 3-D, and panoramic images for Macintosh and Windows. QuickTime is also the leading video production platform for both Windows and Macintosh. Most multimedia on computers begins with or involves QuickTime. Accordingly, the QuickTime technology is a natural for high-quality audio and video playback over the Web. Similar to Windows Media, QuickTime does not charge licensing fees for the number of simultaneous streams served. QuickTime can be streamed from the Mac OS X Server, the

Darwin Streaming Media Server, and RealNetworks' RealServer 8.0.

The latest version, QuickTime 4, features many enhancements including:

- Smaller "component" codec architecture so that the initial download is as low as 1.7 MB. Additional codecs are transparently downloaded in the background when required for a specific media element on a page.
- Support for an increased number of formats, including MP3, Flash, MIDI, and almost every audio, video, animation, 3-D, and virtual reality format available.
- Improved codecs.
- True RTSP streaming when used in conjunction with the Mac OS X Server.

One of the keys to the success of the QuickTime technology and plug-in is that it can handle all types of media elements. For those of you trying to design for the greatest number of users and the least number of plug-ins, this can be a significant benefit.

In addition to playing MP3 content, QuickTime supports Timecode tracks as well as MIDI standards, including the Roland Sound Canvas and GS format extensions. QuickTime also supports key standards for web streaming, including HTTP, RTP, and RTSP. Plus, QuickTime supports every major file format for images, including JPEG, BMP, PICT, PNG, and GIF. QuickTime also features built-in support for digital video, including MiniDV, DVCPPro, and DVCam camcorder formats, as well as support for AVI, AVR, MPEG-1, and OpenDML.

Finally, the newly designed interface is attractive and user friendly. In addition to the traditional controls you'd expect to find on a television--like volume controls and pause and play buttons--the QuickTime Player gives you enhanced controls for online movie playback. The QuickTime Player's LCD section includes a time display, a time slider that shows you the length of the file being played, and a chapter marker. You can switch chapters on the fly even at the beginning of a video stream.

## **Flash and Director Shockwave**

Macromedia Flash is the solution for full-scale, high-impact web multimedia with short sound effects and loops. Flash's bandwidth- friendly vector animation is ideally suited for web content delivery. Flash encodes embedded soundtracks in MP3 format that allows for better streaming and higher quality audio playback.

Flash is also tightly integrated with RealMedia. You can combine a Flash animation with

a RealAudio soundtrack using the RealDeveloper tools to encode a RealFlash presentation. RealFlash allows linear playback from within the RealMedia architecture taking advantage of RealMedia's advanced bandwidth negotiation for streaming audio and video and Flash's streamlined vector graphics for interactive animation.

Director Shockwave is the format of choice for building complex "CD-ROM-like" interactive web presentations and games that utilize Macromedia's powerful Lingo scripting language. Originally designed for full-scale development of interactive CD-ROM content, Director has been retooled to export highly advanced interactive Shockwave presentations for the Web.

Although Macromedia continues to integrate Flash's vector technology into Director and some of Director's advanced programming features into Flash, Director still stands apart in its support for the Lingo script. To preserve its highly compact plug-in file sizes and ease of use, Flash does not incorporate Lingo. Lingo is a powerful scripting language that enables developers to create and customize much more interesting interactive media such as complex strategy games, compelling music videos, and educational tools.

## **Beatnik's Rich Music Format (RMF)**

Beatnik's Rich Music Format (RMF) is an HTML-based format that utilizes common scripting languages such as JavaScript to sync sophisticated interactive soundtracks that combine MIDI sounds and short audio samples to web content. Beatnik allows you to create full-scale, multilayered, interactive soundtracks and compositions that transform and change with user actions. Beatnik presentations sound excellent and download fast. And Beatnik can be incorporated into a web page along with other technologies such as commerce engines and backend databases.

Beatnik has a few distinct advantages over technologies such as Shockwave and Flash. For one, it uses MIDI, a highly compact language for scoring music, to play back audio from a dedicated synthesizer engine such as the Beatnik player. With the same file size (15 to 30 KB) as a two-second Flash audio loop, Beatnik can transmit a great-sounding MIDI score several minutes in length. But Beatnik is much more than MIDI--it also supports the delivery and playback of short customizable digital audio samples, making it far richer than MIDI playback alone.

The downside to Beatnik is that it has a steep learning curve and takes a considerable amount of time to debug to ensure smooth playback. Unlike Flash, Beatnik relies on the Beatnik plug-in, as well as a scripting language like JavaScript to control audio playback and synchronization. Beatnik will likely become more stable and reliable as the technology is refined and as more authoring tools such as Dreamweaver and NetObjects begin to include built-in JavaScript support.

## MP3

MP3 has gained huge popularity as an encoding format because of its great sound quality. For radio-style broadcasts, professionals unanimously agree that it is the best-sounding format. MP3 is most commonly used for easily and efficiently uploading and downloading music files to the Web. MP3 is especially popular among downloadable music enthusiasts because it preserves audio quality while creating file sizes that are up to 12 times smaller than uncompressed WAV or AIFF audio files. MP3 is also quickly becoming the preferred format for streaming music as well, even though it is more complicated than setting up a RealMedia Server.

MP3 is derived from the group known as MPEG (Moving Pictures Experts Group). The members of MPEG are responsible for establishing standards for digital encoding of moving pictures and audio.

Unlike Liquid Audio, MP3 is not a proprietary end-to-end music delivery system. This distinction is important since companies concerned about copyright protection and secure delivery may decide to use the Liquid Music System for music distribution instead of merely posting MP3 files on a web page. On the other hand, the fact that MP3 is an accessible standard means it has the advantage of widespread industry support and compatibility with many applications and media players, including RealPlayer G2, Beatnik, Shockwave, QuickTime 4, and Windows Media.

## Liquid Audio

Liquid Audio provides a complete end-to-end solution for secure music delivery over the Internet. Unlike Flash or RealAudio, Liquid Audio is less of a sound design format for adding audio to your web site than it is a professional utility for music sales and distribution. Accordingly, if you want to sell digitized music files over the web, Liquid Audio is the clear choice. You can purchase a starter package for less than \$1,000. If you just want to broadcast audio so listeners can preview your music, you may wish to use a less expensive option such as RealAudio or MP3.

The Liquid Music System consists of four core products: Liquifier Pro, Liquid Server, Liquid Player, and Liquid Express. Every component of the Liquid Music System has been designed specifically for electronic music distribution. Here is what each component lets you do:

### *Liquifier Pro*

Liquifier Pro is an encoder that allows you to prepare and publish CD-quality, copy-protected music for purchase and delivery via the Internet. The Liquifier Pro

includes DSP functions such as sample-rate conversion, four-band parametric EQ, and dynamics processing, and it provides the capability to include lyrics, credits, and artwork--all in one audio file.

What distinguishes the Liquifier Pro from other encoders is its powerful watermarking and anti-piracy protections. Liquid Audio watermarking inaudibly embeds digital data, which identifies authentic copies of the music into the audio file. Liquid Audio employs multilayer security, which provides data on who owns the music and who bought the music.

### *Liquid Server*

Liquid Server lets you publish and host Liquid Tracks. The Liquid Server also includes an SQL database and can even hook into larger, industry-standard SQL databases, such as those from Informix and Oracle. The flexible design of the server allows you to send dynamic product and promotional information such as sale prices, tour schedules, discounts, and coupons, along with the Liquid Track to be received by the Liquid Player.

### *Liquid Player*

Liquid Player allows you to preview and purchase CD-quality Liquid Tracks on your Macintosh or Windows PC. The Liquid Player is software that lets you preview or purchase CD-quality music from the Internet. It also allows you to see album graphics, lyrics, liner notes, and promotions while listening, as well as easily record a standard "Red Book" audio CD that is playable on any home, car, or portable stereo system.

### *Liquid Express*

Liquid Express is a software package specifically designed for audio professionals in film, radio, television, music, and advertising that allows for the secure real-time preview, approval, delivery, and archiving of broadcast-quality audio.

Liquid Audio now also supports secure MP3 delivery through its watermarking and file security technology. This helps ensure that appropriate copyright and security information will adhere to MP3 files distributed over the Internet, providing the first step to some form of copyright standardization and unified structure to MP3 delivery.

## **MIDI**

Although MIDI is not a streaming format, it downloads so quickly and is so widely used that we decided to include it in this list. If you are looking for an easy, low-cost solution for adding a little theme music or a button rollover sound to your web site, but you don't want the long waits associated with downloading digitized audio clips, MIDI may be a

great option.

MIDI (Musical Instrument Digital Interface) is a super-compact musical language that transmits instructions such as pitch, volume, and note duration to MIDI-compatible sound cards and synthesizers. Since MIDI is a text-based musical scoring language, it downloads super-fast and is ideally suited for HTTP delivery.

The downside is that MIDI is not sound itself; rather, it is the coded representation or score of how the sound should be reproduced by the user's MIDI sound engine. Many browsers and computer systems feature different MIDI sound engines that greatly vary in quality and instrument playback style. This variation makes it difficult for developers to predict what an end user is going to hear.

## Selecting the right format

Each format discussed in this chapter has advantages and disadvantages depending on the requirements of your project. There is no single format appropriate for every situation. To determine which format is best for you, first identify your needs, then select the format that best suits those needs. There are huge differences in server requirements for broadcasting CD-quality music to a limited audience versus wide-scale broadcasting to a large audience with diverse bandwidth capacity. Similarly, the differences between authoring and delivering interactive content such as a game or product demo versus encoding and broadcasting a video file are completely different.

RealAudio, MP3, and Flash are familiar names, but a host of alternative formats, including Windows Media, RMF, and Liquid Audio might better suit your needs. Let's take a look at the factors that will determine the most appropriate format for you.

## Interactive sound design capabilities

Before you look at browser compatibility, cost, audio fidelity, and server performance, you will need to determine whether you need a format that supports interactive presentations or one that supports continuous playback of audio and video files. Several formats such as Flash, Shockwave, and Beatnik are designed for rich interactive media content such as games, educational material, product demos, and promotional pieces where instantaneous feedback via sound effects are essential.

In contrast, formats such as RealMedia, MP3, Windows Media, and QuickTime are primarily designed for continuous playback of audio and video files where server-side bandwidth negotiation and management are key. When they do support interactivity, it is usually in a more limited form such as slide shows and synchronized sound with video or text.

First, determine whether you want a format for delivering interactive content or simply a format for encoding and broadcasting audio and video files, then let the following criteria guide your final decision.

## **Browser compatibility**

Let's face it, if people do not have the plug-in or technology to view or listen to your content, it is much harder to get your message across. This does not mean that you have to select a format that has 100% acceptance, but you will need to assess how tech-savvy your audience is and what format is going to be the most widespread among your target audience. If you are targeting a tech-savvy audience, they will be more likely to download the newest version of the plug-in if they do not have it installed. Do not count on a less technical audience to successfully download and install new technologies.

## **Cost for streaming audio**

To add streaming to your web site, you may need to purchase one or more of the following:

- Encoding software to convert your raw media files into the appropriate format for web delivery
- Dedicated server software to stream your encoded media files
- Hardware to install your server on (this may include several systems for redundancy and scalability)
- Bandwidth to transmit data from your server over the Internet

The cost of streaming can range from free to hundreds of thousands of dollars depending on how many of the above items you will need to purchase. For example, if you are already running an NT server and have a dedicated T1 line, you can use Windows Media for no extra cost. Some vendors provide free introductory-level streaming solutions such as Real Networks. If you have administrative access privilege to your web server, you can install the free Basic RealServer G2. Alternatively, if you are a multimedia producer and own a copy of Macromedia Director or Flash, you can export Shockwave files and stream them from your regular web server for free.

On the other end of the spectrum, if you are a major Internet portal running a hot Sun machine with thousands of simultaneous listeners, you are going to need a healthy budget

for equipment and bandwidth and, if you are using RealMedia, the appropriate number of streaming licenses. Keep in mind that you may not need to spend as much on streaming licenses as you may have thought. A 60-stream license can go a long way. If your average user listens to your audio for five minutes or less, you can deliver 17,280 streams per day with a 60-stream license.

## **Learning curve and documentation support**

As is the case with limited budgets, everyone has a tolerance level for learning new technologies. Keep in mind that some formats provide much more documentation and software tools for getting started. RealMedia, for example, has outstanding documentation and software support, including sophisticated tools for automatic server configuration. In contrast, other formats such as MP3 and MIDI are not all-in-one proprietary streaming solutions but are merely standards for audio compression or musical notation and thus do not offer a single source for documentation and support.

Besides documentation and support, the real hurdle depends on the scale of your streaming needs. If you merely want to broadcast the annual company report to a few hundred nationwide sales representatives, streaming audio is a much simpler affair versus competing with Spinner.com to become the king of Internet radio. The difference in the infrastructure required for streaming to a few hundred listeners per day versus tens of thousands is night and day. If you are broadcasting to a huge audience with a scalable robust system, the learning curve is going to be much steeper than setting up a free Basic RealServer or throwing some audio files up on your HTTP web server. Large-scale professional broadcasting requires advanced configurations and logistics, such as multicasting with multiple servers and backup systems in place for redundancy.

## **Audio fidelity and compression**

Audio fidelity for the end listener is determined by the quality and specific setting of the codec used for audio compression and decompression. Better compression algorithms, such as MP3, result in higher fidelity audio playback over the same bandwidth connection. Audio fidelity is also determined by the target file size and bandwidth settings you are using when encoding the sound file. A larger target file size requires less compression and audio degradation but more end-listener bandwidth capacity.

Server performance--the ability of the server to detect and send the appropriate stream to the end listener--is often just as important a factor in producing an overall quality listening experience as the codec.

## **Low bandwidth performance overall**

Nobody likes to scrimp on quality or excitement, but if you have to tailor your media to fit the lowest common denominator of your audience, you will have to make some tough choices. Some formats, such as RealMedia, excel in bandwidth and browser compatibility. Other formats, such as Shockwave or Flash, work better in high-bandwidth 56 Kbps and DSL environments and provide little to no support for server-side bandwidth negotiation.

There are two factors to consider when selecting a format for low-bandwidth environments: the inherent ability of the format to provide compelling media with small file sizes, and the server-side technology to manage the delivery of media when constrained by low or fluctuating bandwidths.

Beatnik, for example, packs a huge punch of interactive excitement in an extremely compact file size because it utilizes MIDI. The use of bandwidth-friendly MIDI technology gives Beatnik an inherent advantage over Shockwave or RealMedia. On the other hand, RealMedia provides better server-side support for ensuring that files get delivered and do not drop out, regardless of the bandwidth of the end user.

## Server performance and software quality

Thinking big? For those of you who need to stream audio and video content to thousands of simultaneous listeners on the scale of a CNN, NPR, or C|net, you will need a format that provides powerful server-side features and tools. And if you plan on broadcasting live events, you will need a real-time encoding and streaming system that runs on a dedicated web server.

RealMedia and Windows Media are the leading technologies for large-scale broadcasting, with SHOUTcast (MP3) and QuickTime close runners-up. The RealServer and Windows Media Server provide bandwidth negotiation that ensures smooth audio playback for the end listener and prevents annoying drop-outs when bandwidth fluctuates.

Beyond the actual server software you choose to install, whether it's RealMedia, Windows Media, SHOUTcast (MP3), or QuickTime, streaming to a large audience is just as much or more about the hardware and bandwidth as the format you choose. Large-scale broadcasting requires multiple systems, servers, and huge bandwidth connections. That's why many companies outsource their media broadcasting to companies like Broadcast.com or Network24.com. For a further analysis of the characteristics of each format, refer to Appendix B, *Audio Format Comparison*. It contains a chart that will help you select the appropriate format.

## Summary

With so many technology options available for delivering media over the Web, it can be a challenging process to try to select the appropriate format for your multimedia web pages. This chapter attempts to point you in the right direction and provides you with a thumbnail view of the many options available.

It's time now to take an in-depth look at the most popular audio solutions available on the Web. We will begin with RealAudio.

**Back to: [Designing Web Audio](#)**

---

[O'Reilly Home](#) | [O'Reilly Bookstores](#) | [How to Order](#) | [O'Reilly Contacts International](#) | [About O'Reilly](#) | [Affiliated Companies](#)

© 2001, O'Reilly & Associates, Inc.  
[webmaster@oreilly.com](mailto:webmaster@oreilly.com)