

## Chapter 8: Overview of Streaming Audio and Video

Since the explosion of interest in the Internet in 1993, people have experimented with transmitting sound and video over the Net. As we saw in Chapter 7, for the most part, this was a disappointing experience, because of the time it takes to transfer an entire multimedia file over slow links. An audio file might take more real time to download than the length of the clip being played—that is, you might spend 10 or 30 minutes downloading an audio clip whose elapsed playback time might be only two minutes. Video, which carries much more information than audio, entailed even longer download times, just to experience a 1/8 screen, slow-frame-rate, blurry movie.

The advent of streaming media changed all that. Streaming media uses an age-old concept *buffering* to make viable the playback of multimedia content *while it is being downloaded*. A buffer holds a reservoir of content sufficiently large to smooth out the bumps in playback that may be caused by momentary server sluggishness or network congestion.

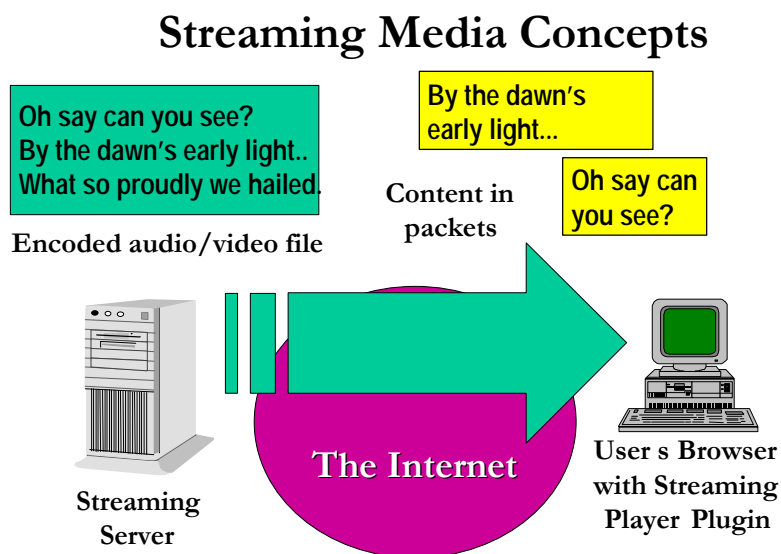
If you've ever used an audio CD player designed for joggers, you've taken advantage of the same basic concept. When you hit a rough spot, the laser that reads the data pits on the CD may skip out of place, which normally would mean you would notice an audible interruption in playback. A memory buffer holds enough seconds of sound for the player to continue the music uninterrupted until the laser can find the right track, and the CD can rotate to the right sector, so as to pick up the data stream. You, the listener, never realize that all of this has happened. Similarly, buffering allows Internet streaming media to (usually) maintain continuous playing of music or speech despite the occasional burp in network delivery.

Streaming media combines this concept of buffered real-time playback with *compression* to make viable what once might have been considered impossible—delivering to hundreds or even thousands of simultaneous listeners. Each of these listeners has his or her own Internet connection from remote server to local desktop, delivering audio in AM radio, or even FM, quality. The story is not quite so happy when it comes to video, because there is so very much information in full-frame, full-motion video content. Nonetheless, real-time streaming of video has also improved greatly as new codecs are designed with better and better compression. Internet streaming video is nowhere near the quality of conventional television, but it is serviceable for some applications, and it will continue to improve.

Your basic steps in sending out content via streaming are:

- Create or obtain content. The content might be a recording you make, or it might be produced as part of a live event. (We'll discuss live events further later in this chapter.)
- Encode the content into the special streaming format
- Use a streaming server to send the content to your listeners.

The following diagram depicts these basic concepts:



Here, we've encoded an audio rendition of the Star Spangled Banner. Playback occurs in this fashion:

- After it was digitized, the entire audio file was placed on a streaming server to await requests for playback. A streaming server is simply a server that has specialized streaming software installed—for instance, the Real System Server from Real Networks, Inc. The content has been specially encoded and placed into a file in the streaming server's file hierarchy.
- The server waits for a request from the user for a particular streaming document. When a user clicks on the URL for such a document, the user's browser sends a request to the streaming server.
- The streaming server finds the relevant content and prepares to send the file over the Internet. As the file transmission begins, the contents are broken into packets; each packet is sent as soon as it is prepared.
- The user's browser has had a streaming media plugin installed, such as the Real Player. The plugin places each packet into its buffer as it arrives, and, when the buffer is sufficiently full, the plugin starts playing the content.
- Further packets continue to arrive. Thus, the buffer is being filled and emptied simultaneously, as playback continues—usually uninterrupted.
- In the event of severe network congestion, playback may pause. In this case, the player the user will experience a pause in the playback while the player attempts to refill the buffer.

### Vendors of Streaming Solutions

There is one dominant vendor of streaming technology, with other players on the horizon:

- Real Networks Inc is the pioneer in streaming technology, and continues with 90% of market share as of this writing. Formerly Progressive Networks, Real created first RealAudio, then RealVideo, formats, and now markets the RealSystem with these and other formats.
- Apple was a pioneer in non-streaming digital video on personal computers with its QuickTime format. With the release of QuickTime version 4 in 1999, Apple began touting QuickTime as a streaming format.
- Microsoft, once a partner with Real Networks, is now advocating its own streaming architecture and its own products such as NetShow and NetMeeting.

As of this writing, your choice of which technology to embrace will be fairly basic: If you select the RealSystem, you will use the format that is most likely to have players installed on your user s desktops. This is because Real Networks was the only streaming media vendor to succeed in getting its plugin bundled with browsers and with the Windows 98 operating system. If you choose another system, your users are more likely to have to install a player plugin before they can play back your content. Thus, you may want to select Real, unless you see features in competing systems that offer a compelling advantage unique to your content. Over time, Microsoft and Apple and perhaps new contenders may well break the dominance Real has as of this writing.

The rest of this chapter will give an overview of the Real System tools and techniques. The general concepts will apply to other systems. Real Networks provides excellent, thorough documentation, which you will want to consult and read in detail when you're ready to launch a project. In particular, look for these two manuals at [www.real.com](http://www.real.com):

- *RealSystem G2 Production Guide*
- *RealServer Administration Guide*

Each of these manuals is over 200 pages in length. They are offered in Adobe Acrobat format. You will probably want to print the content and place it in a binder for offline reading.

Note that the Real software has evolved dramatically over time. During this rapid evolution, content providers and users of the Real System have experienced frustrating compatibility problems. With Real System G2, these problems have largely been ameliorated, and the system has reached a reasonable level of maturity.

### RealSystem Media Types

The RealSystem offers several media types:

- Audio: Suitable for speeches, oral history, and music delivery. Quality can be surprisingly good even over a 28.8 kilobit/second modem.
- Video: Video with audio in synchronization. High-quality video requires a tremendous amount of bandwidth, so over dialup modems or even over faster links such as DSL or cable modems, it is typical to have video in a small window moving at a slow frame rate and with relatively little detail seen on screen. For talking head applications this conveys a sense of the personality of the speaker, but does not fully convey changes in facial expressions, gestures, etc.
- RealPix: This is a format for streaming delivery of photographic slide shows. Each frame is a still photograph, which may be on screen for many seconds; audio is coordinated so that narration or sound events match the time when a new still is downloaded. The effect is somewhat similar to some multimedia slide shows in museums.
- RealText: This format streams textual information and is especially useful in training applications
- RealFlash: This format marries Real with Macromedia's Shockwave Flash format, so that efficient and impressive vector graphics can be sent in a streaming fashion.

There are definite applications for all of these media types, but most content providers have concentrated thus far on the audio and video types. RealPix could be a very promising format to consider for specialized applications, such as a slide show depicting community history through old photographs.

### RealSystem G2

RealSystem G2 offers many important advantages over previous versions of the Real software, versions 1 through 5:

- G2 introduced the RealText and RealPix media types.
- G2 servers are capable of streaming other formats such as WAV and AVI directly, without the need to *encode* content into Real's proprietary format. However, the vendor claims that the most effective streaming occurs only when you do use their format.
- Real introduced SureStream technology, which is a way for you as a content provider to create a single file which includes your content encoded at multiple bandwidths. Prior to this innovation, many content providers would offer a set of hyperlinks for a single document, e.g. one to content encoded for 28.8 kilobits / second, another link for the same content encoded for 128 kilobits / second.
- With SureStream comes automated bandwidth negotiation, so that the user's player and the streaming server can determine the highest bandwidth the network connection can handle dynamically.

Although G2 offers these many important advantages, one problem is that users with older versions of the RealPlayer will not be able to benefit. This is compounded by the fact that Windows 98 came with RealPlayer version 4 installed on the system. You will need to evaluate whether to use G2 features and formats; i.e. will this cause your users to have to download and install a new plugin, and will they be willing to do so?

## Tuning Your Streaming Content to Your Users' Bandwidth

There is a definite tradeoff between the resolution, or bit rate, at which you encode your streaming content, and the quality observed by your users. The higher the bit rate, the better the audio sounds, and the better the video appears.

However, not everyone has high speed network connections. Because you are likely to have users connecting at a variety of speeds, you may want to encode content at more than one bandwidth rate. Let's say most of your users connect over dialup, and some users connect over ISDN. You may want to encode your content at two bit rates - say at about 20 kilobits per second for dialup (with video at a 160 X 120 frame size), and say 100 kilobits / second for ISDN.

Note that you can't encode at a rate equal to the modem speed, because you may not achieve that speed at a sustained pace. Note also that rated bandwidth speeds may overstate true capacity. Many users of 56K dialup modems report that they seldom if ever connect at that rate.

If you encode your data at a higher bit rate than a given user's connection allows, that user will experience frequent pauses in playback, or may be unable to ever play the content at all. This is very frustrating to users. Although you cannot control the vagaries of network congestion on the global Internet, if you do offer your content at a slow enough bit rate for the lowest common denominator, the more likely you are to have satisfied customers.

Traditionally, in order to meet the bandwidth capabilities of a wide variety of users, publishers of streaming content have chosen to encode content at a variety of bandwidths, and offer hyperlinks to each title under each of the different bandwidths. For instance, you might offer content for 28.8 kilobit/second users, for ISDN users (up to 128 kilobits /second) and for Local Area Network users (up to 10 megabits / second). Each of these different speeds, targeted for users with corresponding connection speeds, required a separate, visible hyperlink on your Web site. The users had to know how fast their connection was, and choose among all the links offered.

The RealSystem G2 provides a mechanism called SureStream technology to obviate the need for offering multiple bandwidths. In a nutshell, this scheme stores multiple versions of the content, encoded at every desired bandwidth, in a single file. When the user requests the content, the RealPlayer and RealServer negotiate the appropriate bandwidth based on the user's capabilities. In fact, if network congestion causes interruptions in delivery, the rate can be re-negotiated at a lower value, allowing the user to continue to partake of the content with minimal interruption.

While the SureStream scheme is a great idea in theory, one problem is that previous versions of Real products do not support this mechanism. Because Windows 98 shipped with Real Version 4, users cannot participate in SureStream unless they upgrade their RealPlayers. Therefore, many content providers continue to manually offer separate links to separately

encoded files. (In fact, for that reason, the Toolkit video content is presented as separate files.)

### Components of the RealSystem

The RealSystem provides these tools:

- The RealSystem Authoring Kit. This includes the Encoder, which converts content into the Real format. You create your content in a format you like—say AVI or Quicktime for video, or WAV or AIFF for audio. The Encoder transforms the content into Real's streaming format.
- The RealSystem Server. This is the specialized server that delivers content in real time to your users. Servers are available for Windows NT and for Unix.

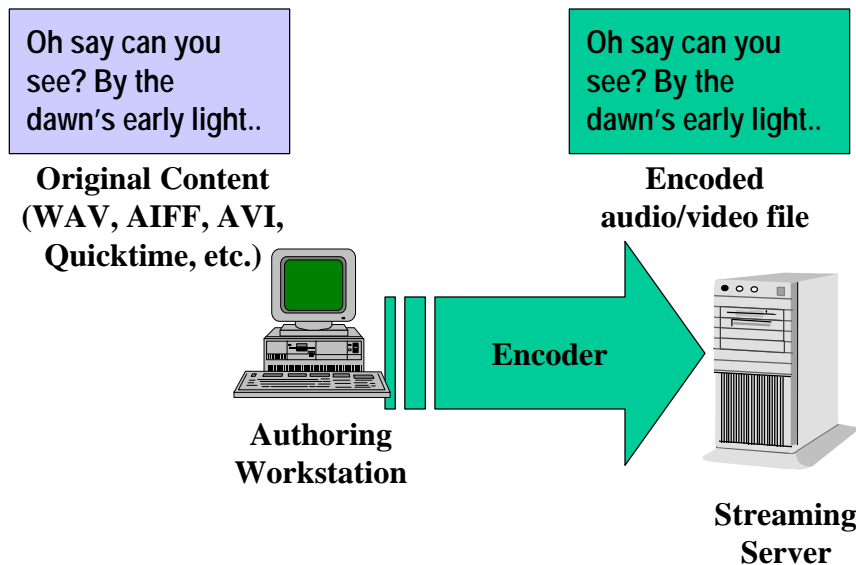
Real offers basic and Plus and Pro versions of these tools. You may want to evaluate streaming media using the basic versions, and graduate to the for-pay versions, which offer more features, when you decide to commit to streaming media as a methodology.

### Preparing Your Content

You will want to prepare your content using formats that are familiar to you and tools you have available. For audio applications, use the advice in Chapter 7 to determine how you will prepare audio content. Video editing demands quite a bit more in the way of computer resources than audio editing, but desktop video editing for streaming purposes is certainly viable today. You will need a video-ready computer, or a video capture card such as the ATI All-in-Wonder or a TrueVision capture card. You will also need editing software such as Adobe Premiere. Alternatively, you may wish to rent or gain access to a professional editing device, such as a Media 100 or an Avid, used by TV stations, video production houses, and available at some schools.

If you are capturing original content, you may do need a great deal of editing in order to make the material satisfactory. You may want to eliminate unwanted passages, or even re-order and significantly reduce the content as would be done for a documentary. Or, for some material, you may decide to make all the raw content available, a la C-Span. Edited material appeals more to some audiences; others like the cinema verite approach.

## Encoding



This diagram depicts encoding of content that you have digitized into one of several popular audio or video formats. It is also possible to encode content into the Real format in real time. You would do real-time encoding for a live event, or you can encode straight from an external source such as video tape or audio tape into the Real format.

Note that the content, once encoded, is only usable by the RealSystem. A Real-encoded file is not meaningful outside the context of the RealPlayer, and there are no tools to translate from the Real encoding scheme back to any other audio or video format. Therefore, if you began with content in some other format, you probably will want to archive your source material even after encoding, in case you want to do further editing in the future.

### License Considerations

Real Networks offers free tools for preparing and playing content. It sells licenses for serving the content, as well as selling advanced content preparation and playback tools. Although normally Real charges for streaming server technology, you may want to take advantage of exceptions to this:

- Real Networks offers a Basic server with up to 25 simultaneous streams for free. A single entity can install one such server.
- School, library, governmental, and other nonprofits may be able to serve for free. Consult [www.real.com](http://www.real.com) for current license information.

A single stream is consumed for each user who connects to your streaming server for content at a time. If 1000 people are connected, you need a license for 1000 streams.

### Real Time Events

You may consider offering over-the-Internet playback of a real time event, such as a speech or a concert. As with all your technology choices, you can do this yourself, or you can have someone else host the event. There may be many partners to consider as your hosting service:

- Many local radio stations now broadcast via Real in addition to over the air. If you convince a radio station to cover the event, you get the streaming distribution for free, and you pick up radio listeners as well.
- Some Internet Service Providers and other Internet technology companies specialize in handling live events.
- For major events, you may wish to choose a national provider. Real Networks operates Real Broadcast Networks, which hosts content for an hourly fee. Broadcast.com, now a subsidiary of Yahoo, also offers such services.

If you use a service to host your real time event, you will need to decide whether you are encoding the content at the site of the event, or at a spot closer to your service provider. For instance, for an audio-only event, let's say you're having a local service provider host the event. You could put a server with the real time Encoder at the site of the event, and send a digital stream over an Internet connection back to your ISP. A better solution might be to send content back using specialized hardware such as a Zephyr box, a proprietary device tuned to send high-quality audio over an ISDN link. By using specialized hardware and a dedicated link, you ensure the content makes it to your ISP intact with no network congestion-induced glitches. You also save the hassle of setting up a PC in the field. In this scenario, your content will be encoded at some remote location, such as the computer room at your office or your ISP.

Real time streaming may be appealing at first glance, but do consider the downside: for a single event, it is unlikely that you will be able to reach a large audience. In the unlikely event that you do reach a large audience, you will need a license for a large number of streams, and this can be expensive.

In the community networking context, it may be more appropriate to think of archival material than real-time events. For instance, a collection of oral history recordings made available via a streaming server could be played any time by any listener anywhere on the planet. Real time, by contrast, assumes that every listener is sitting at his or her computer, connected to the Internet in general and your event in particular, all at once. Although any live event over the Internet is likely to draw some listeners due to the relative newness of this medium, it is very difficult in practice to draw an audience of more than a handful of people. A permanent archive will eventually serve more users, because they can plug in to your content on their own schedules.

If you do choose to try a real-time event, note that the streaming server is capable of capturing an archival copy of an entire event on disk as the netcast occurs. This will allow you to serve both real-time and after-the-fact participants.

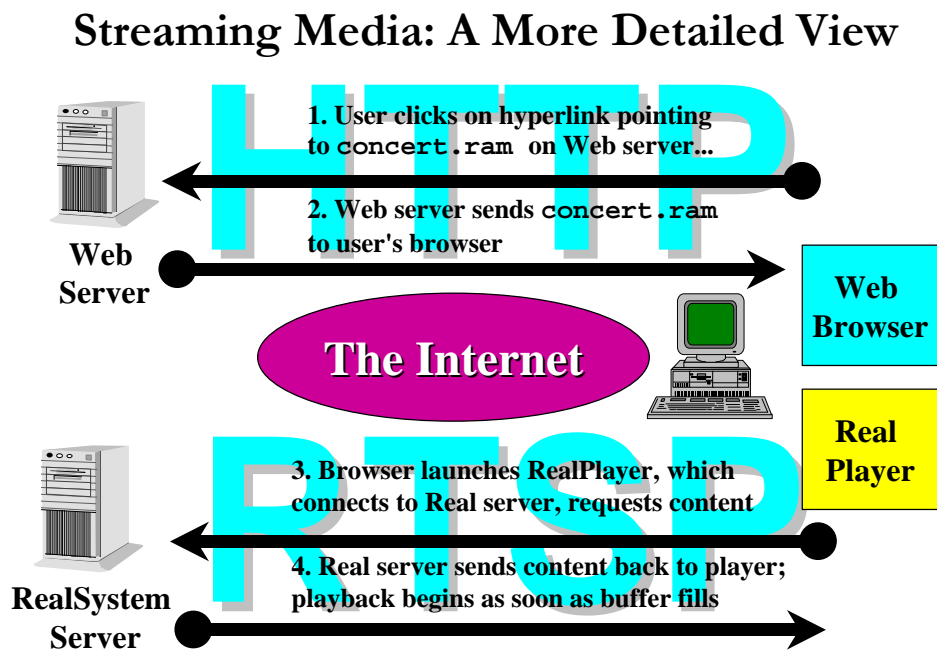


## Understanding the Protocols

Our description of how streaming works at the beginning of this chapter simplified things a bit. As you prepare to serve Real content, you will need to understand how a user is handed off from a Web server to a Real server in order to deliver content.

As users surf the Web, opening URLs and clicking on hyperlinks, content is delivered from a Web server using HTTP—the Hypertext Transfer Protocol. This is the protocol spoken by all Web servers. When you download content in a streaming fashion, you actually will use a different protocol, known as RTSP, or Real Time Streaming Protocol. (An older protocol called Progressive Networks Media, or PNM, is used by older Real servers; RTSP is an Internet standard and is used by default by the G2 server.)

In effect, when a user clicks on a link to a piece of Real content, he or she in fact clicks on a link to a .ram file. This file is actually a tiny little pointer to the actual content. The following diagram shows how the process actually works:



The user's initial encounter with Real content will occur over the Web, using the Web protocol of HTTP. Once the .ram file is downloaded, the user's browser will launch the RealPlayer to handle the file. The RealPlayer will read the .ram file, figure out what server and file name it must request, and open an RTSP connection to that server. The encoded Real content will be sent from the streaming server to the RealPlayer and played during download as discussed earlier. The .ram file is tiny: it contains a single line of information, which is a URL that looks something like this:

```
rtsp://streaming.smallville.org/events/gettysburg.rm
```

While this URL may look a little different than the URL examples we've seen previously, it is actually quite consistent. The URL begins with RTSP because that is the protocol used to fetch the content, instead of the familiar Web protocol of HTTP. The rest of the URL follows the familiar form: our streaming server has a host name of `streaming.smallville.org`, and the path to our particular RealMedia content is: `/events/gettysburg.rm`

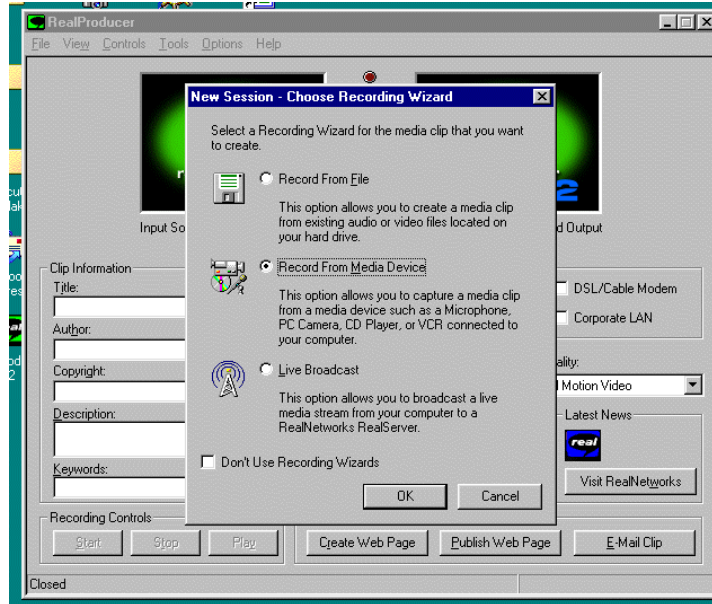
This scheme requires the content provider to create a `.ram` file on the Web server for each piece of Real content to be served, and a `.rm` (RealMedia) file on the Real, or streaming server to house the actual encoded content.

Note that in many cases, if not most, the Web server and the streaming server will be the same physical server hardware, operating under a single Internet host address. By default, Web servers are assigned TCP port 80, and Real assumes a different port for its purposes. (TCP ports are an Internet mechanism to allow more than one service to operate on a single server box; see Chapter 9 for more information on this concept. Real uses port 554 by default, but this fact is usually hidden to the user and the server administrator; the player knows what port to use to connect to the server, and, unless the administrator has a reason to depart from the default, the port number need not be specified in the RAM file.) Thus, a single server—for instance, your main Web server at `www.smallville.org`—can house both Web server software and RealServer software, both running in tandem on the same box. In this case, the HTML documents that refer to Real content, the `.ram` files, and the content itself in the form of `.rm` files, all reside on a single physical server.

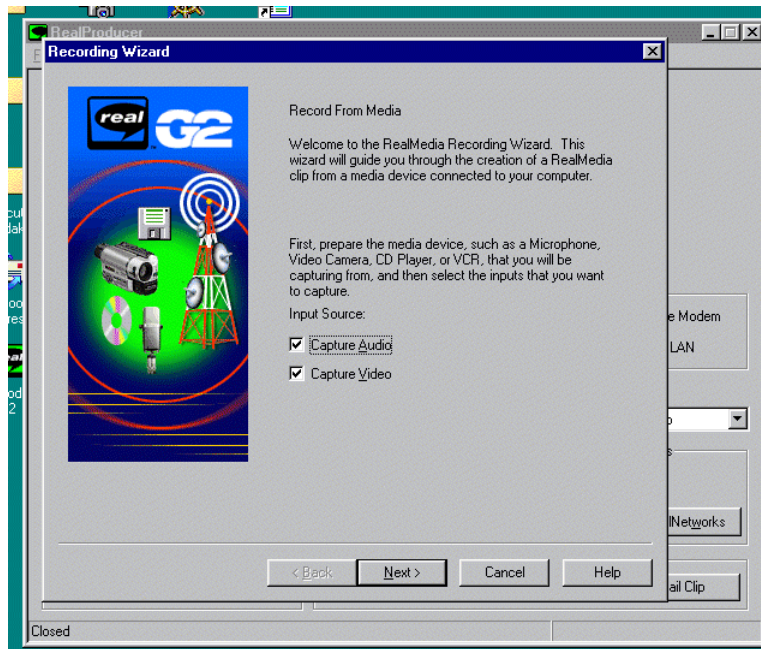
In order to relieve some burden on content providers, RealSystem G2 provides a `ramgen` mechanism to create `.ram` files automatically, eliminating a tedious task. See the Real documentation for details.

### Wizard-Based Content Preparation

RealSystem G2 provides a set of wizards to walk you through the process of preparing content for delivery on your Real server.

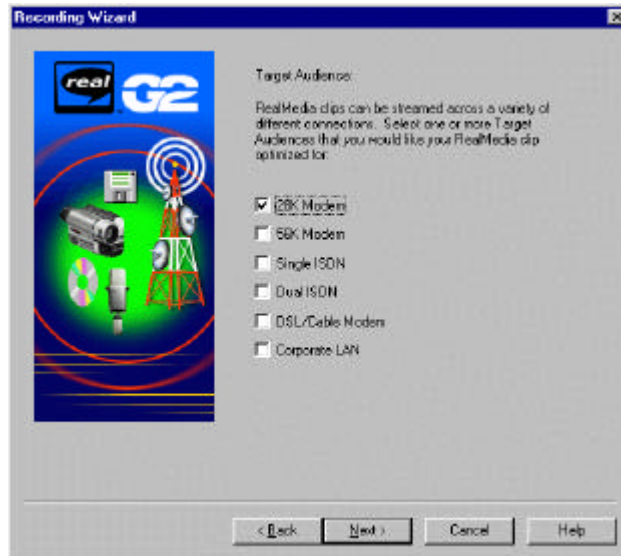


First you tell the wizard what kind of content it will be working with: an existing digital media file (such as in AVI or QuickTime format), a media device such as a tape deck, or a live event broadcast.

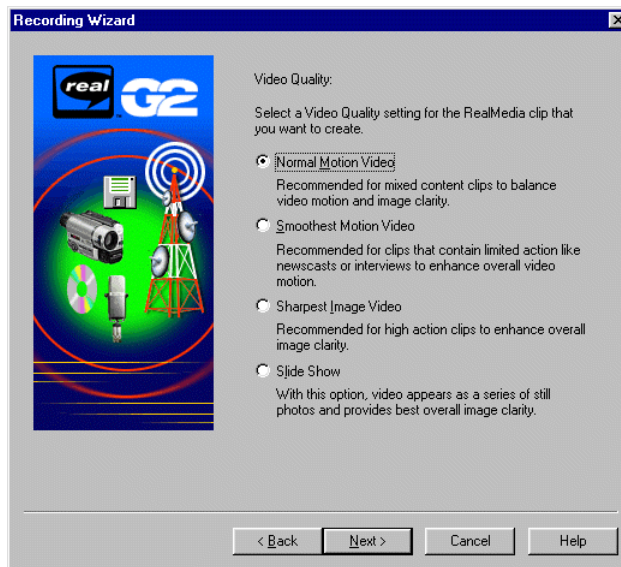


In this case, we're working with video content from a VHS VCR connected to an ATI All-in-Wonder card on a PC. The ATI card accepts video and audio using conventional RCA jacks from the video and audio output plugs on the VCR. If the VCR has S-video, the ATI card can accept that as well. This provides a video signal superior to that offered by the RCA plug. If you are using a video format better than VHS, such as Hi-8 or Mini-DV, you will want to use S-video, which transfers video with higher fidelity than the older RCA-style connection. (Mini-DV camcorders and decks can also output the audio and video digitally

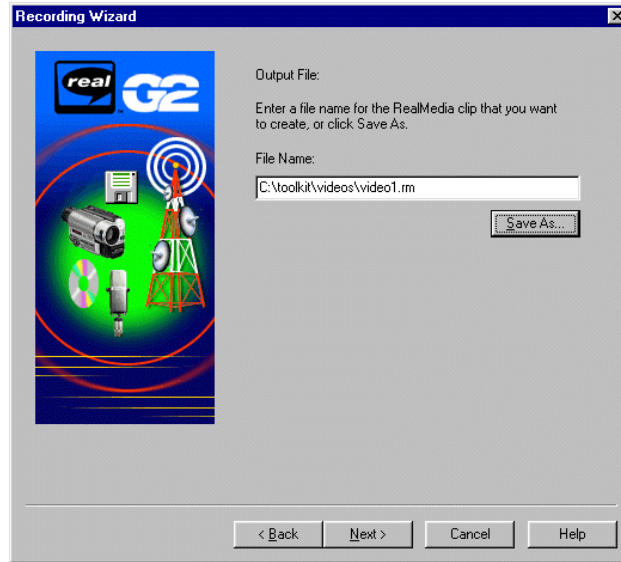
using IEEE-1394 or FireWire connectors, if your PC is equipped with a FireWire port, offering the best fidelity of transfer possible in consumer video.)



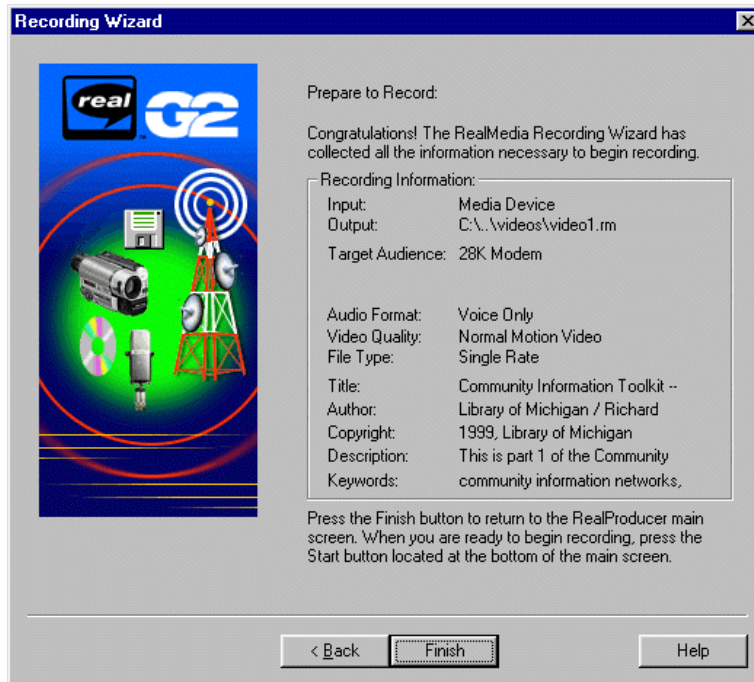
Here the wizard asks what the target audience is, so as to set the encoding rate appropriately.



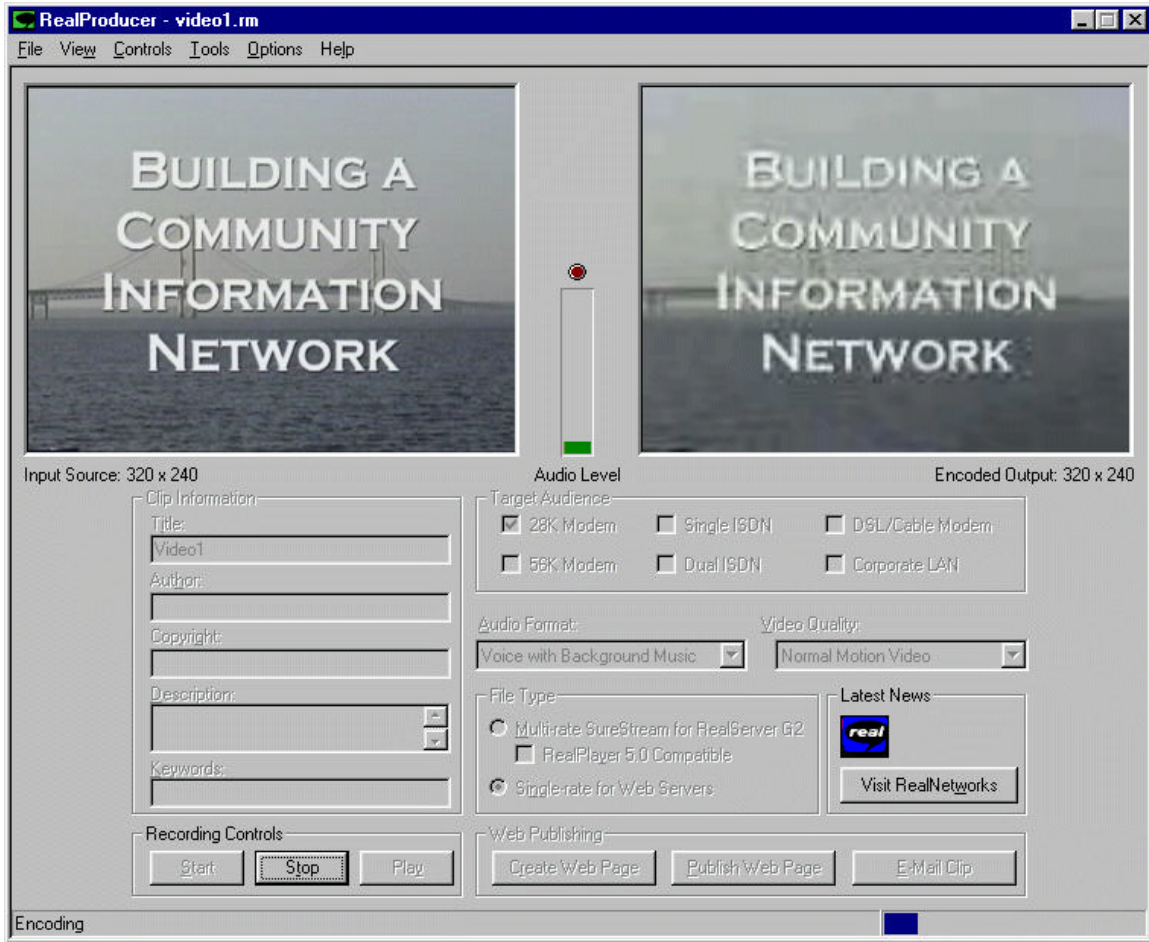
The RealSystem can make tradeoffs in image quality, tuning for the kind of content you have.



Here we tell the wizard what file name to save the content under. Now we're ready to begin recording.



As the wizard advises, a new screen will appear after you click Finish, and you can press Start on that screen when you've got your tape cued for playback.



During the encoding process, the panel on the right will show an approximate view of what the encoded video content will look like, given the constraints of the window size (in this case 320 X 240 pixels) and the encoding rate. Look carefully at the two panels; note that the text and the image of the Mackinac Bridge are somewhat blurrier on the right.

Encoding will continue until you press Stop on the control panel on screen. If you fail to stop encoding at the end of your source material, you will end up with spurious material in your RealMedia file.

You may find that you have better control over the process if you digitize into a format such as AVI, and use any video editing tools you have at your disposal to make necessary edits. For instance, if you have leader or trailer information, you can remove it with the editing software, instead of babysitting the transfer from tape in real time. Note that formats such as AVI consume prodigious amounts of disk space; you'll need gigabytes of disk for minutes of content. (By definition, Real-encoded content is highly-compressed; if you are constrained for disk space, this is an advantage of encoding directly into Real from your media source.)

Once you have prepared encoded content, you can preview it locally on your hard disk using the RealPlayer. When you are ready to move the content to the RealServer, another wizard

automates the process of publishing your page, even generating the necessary .ram file for you, and transferring the files to the RealServer you specify. It adjusts links in the HTML and .ram files to point correctly to the right locations on the RealServer to which you are posting the content.