DELIVERING VIDEO TEACHING COURSES THROUGH THE INTERNET: TECHNICAL ISSUES

M. Crudele[1]

Streaming video technology allows nowadays the delivery of video based courses over a standard TCP/IP network, either Intranet or Internet, even at low bandwidth rates. We discuss the technical issues raised during more than one year of experience in converting videotapes to streaming format and in delivering them to students. We conclude that the quality, performance and cost achieved give any institution the possibility of setting up a video server for storing their own educational videotapes, to make them available at any time from any multimedia computer on the internal or external network.

1. INTRODUCTION

In a previous paper [1] we described the Multimedia Production Process [2] for converting existing videotape based educational material into a format suitable for delivery over standard IP networks, even at low bandwidth (typically 20 Kbps). Since then, due to the advancements in the compression technology and some refinements in the methodology, we have been able to fully test the process, setting up a video server.
Educational software is generally considered to be difficult and costly to implement. Revenues for developers are small, mainly because schools are not investing too much money in software [3]. Designing new software for teaching is also a hard task and takes time. On the contrary, there is a very wide production, still ongoing, of videotape based educational material. Moreover, a typical lesson can be now videorecorded at a very low cost with a good quality, using amateur videocameras.
Our primary goal was to develop a procedure to convert the videotapes for using them in a multimedia laboratory Intranet, so that any student could access the desired one at any time, at his or her own pace. The constraints were to make the procedure cheap, fast and easy, which means affordable for any institution. The project, extended to Internet access, is still going on, with further enhancements. Meanwhile, we have been able to implement and test the procedure, as well as the delivery of a course to a group of students.

2. MATERIALS AND METHODS

2.1 Streaming video overview

Streaming audio and video is a method of transmitting in packets that are played as the system receives them, meaning you don't have to download any file to the hard disk. By streaming multimedia content, you can view lengthy video segments and entire movies in real-time, provided you have the sufficient bandwidth for the encoded stream. In the following sections we emphasise on the main topics related with this technology.

2.1.1 Compression

Compression methods shrink the video signal down to a size that can be transmitted across available network bandwidths. As a point of reference, uncompressed, small size, 30 frame per second video signal can consume over 18 Mbits per second (Mbps). Most streaming video products are based on MPEG, H.263 or wavelet compression methods. Both MPEG and H.263 are known as differential coding methods. Differential coding methods take advantage of the fact that much information remains the same from one frame to the next (e.g., the background). In this method, only the differences require transmission. H.263 was designed for modem-based videoconferencing and performs better at speeds ranging from 28.8 through 256 Kbps.
Wavelets work by coding successive layers of resolution for each video frame, and then playing them back at

---

[1] Medical Informatics Laboratory, “Campus Bio-Medico” University of Rome
e-mail: m.crudele@unicampus.it

the highest level of resolution that the available bandwidth will allow. The advantage of wavelets is that a single video file can be played out over multiple bandwidths. The disadvantage is that wavelets are forced to transmit each individual frame in its entirety. This individual frame transition causes difficulty at low bandwidths in attaining the frame rates necessary to convey motion.

MPEG-1 and MPEG-2 [4] were designed for CD-ROM and commercial broadcast applications, respectively. They offer excellent clarity and motion representation, but they are not suitable for transmission speeds under one Mbps. MPEG-1 needs, at the standard frame size SIF (352x288 for PAL system, 352x240 for NTSC, because of the different fields per second rate, 50 and 60 respectively), a bit rate of 1.5 Mbps, which can be achieved on a 10Mb LAN, or on a T-1 (1.54 Mbps) or PRI ISDN (2 Mbps) connection. However, the number of concurrent streams is necessarily very limited.

The MPEG-4 standard [5] is in its final stage and will provide in 1999 a set of technologies to satisfy the needs of authors, service providers and end users. It uses part of H.263, the file structure of Quicktime [6] and will contain differential coding methods while also incorporating many of the scalability attributes of wavelets. Some features will be: speed change (to implement a "fast forward" function); pitch change (for voice alteration); bitrate and bandwidth scalability (for adapting to the actual bandwidth); error robustness (to avoid or conceal audible distortion caused by transmission errors); audio effects (to achieve functions for mixing, reverberation, spatialization, etc.). The bitrates supported will span between 5 Kbps and 4 Mbps; the formats will be progressive as well as interlaced video; the resolution, from small to TV size.

MPEG-7, due after year 2000, will also define the manner in which audiovisual materials can be coded and classified so they can easily be searched through.

2.1.2 Networking

The main delivery mechanisms are Transmission Control Protocol (TCP) and User Datagram Protocol (UDP). The problem with TCP is that, with its error free capability, it causes retransmission of lost packets, which can produce long pauses, not acceptable for audio or video, especially on live events. The problem is somehow solved by buffering, that is accepting more data before it is actually needed: in live events, there is always a delay of quite a few seconds between reality and playback.

UDP does not retransmit lost or wrong packets, and therefore assures a smooth play. Therefore a system for compensating some of the lost packets must be provided. The techniques are: over-sampling, in which redundant data is sent; interleaving, in which portions of data are spread across multiple packets; filtering, in which received errors are somehow hidden or transformed when the are received.

Bypassing proxies and firewalls, which normally do not allow UDP transport or TCP packets other that for http, is another issue. Almost all available products allow now the HTTP-like mode that uses standard http, even though it loses performance, while eliminating the need of setting up a specific software into the firewall.

2.1.3 Multicasting

Multicasting is a way of sending a live broadcast as a single data stream to only those clients who request the data. Multicasting contrasts with two other modes of transmission:

- Unicasting, which sends a separate, point-to-point data stream to each client that requests it.
- Broadcasting, which sends a single data stream to all clients on a subnet, regardless of whether a client has requested the data.

Multicasting consumes much less bandwidth than either unicasting or broadcasting. Multicasting sends only a single data stream (rather than many copies of the data stream) only to those clients who request the data (rather than all clients on a subnet). Multicast delivery over the Internet is made possible to a limited degree by the Internet Multicast Backbone (Mbone), which is a virtual network consisting of those portions of the Internet that are multicast-enabled. For the most part, however, multicasting is used over intranets.

Multicasting is suitable for delivering live events and for videoconferencing [7], but not for a video server, where each user can choose its own video stream (so called VOD, video on demand).

2.1.4 RealNetworks RealSystem

RealSystem, by Real Networks [8], enables streaming audio and video for reliable transmission under real-life conditions. By using a server software and a client (available for free, as an application as well as a browser's plug-in), the user has complete control over the video clip: pause, move forward and back, stop and restart, or go to specific time points. The encoder is also a part of the system, and it uses a proprietary format, or a fractal scheme. The stream can be enriched by embedding event triggers that can open pages on a standard browser.
The server software we used is a free version (5.0), limited to support 60 concurrent streams. RealServer requires approximately 6 MB of available RAM plus 40 KB RAM for each simultaneous stream. For example, to support 100 simultaneous connections, it requires approximately only 10 MB of available memory. RealServer has also a modest CPU impact. A Pentium 120 MHz processor is sufficient: for 100 concurrent streams , RealServer would consume less than 30% of the CPU cycles. With enough network bandwidth, the same computer can deliver at least 400 simultaneous 28.8 Kbps streams.
At the end of 1998 RealNetworks will release the final version of a new product, called G2, which improves audio and video, using also SmartStream, a sophisticated new transport system that dynamically scales throughput to match available bandwidth. It can use RTSP (Real Time Streaming Protocol, an IETF standard) for client-server transport and SMIL (Synchronized Multimedia Integration Language, a W3C standard) for multimedia layout and integration, which is XML-compliant. The server can be installed on Unix and Windows NT. The hardware requirements are higher than the version 5 ones.

2.1.5 Microsoft NetShow

NetShow is similar to RealSystem, distributed free of charge by Microsoft [9]. It can use various codecs (mainly an early implementation of MPEG-4, plus others from different sources and companies) and has good indexing capabilities. All the other features are similar to RealSystem. Version 3 of the server and the encoder has recently been released. The player is now a standard component of Windows 98, but can be installed on other operating systems. It can also play almost all kinds of video formats. The quality of audio and video has improved noticeably. All the software for producing, delivering and playing video streams is free of charge, because Microsoft is making a deal with Windows NT server, which is the only supported system for the NetShow server.

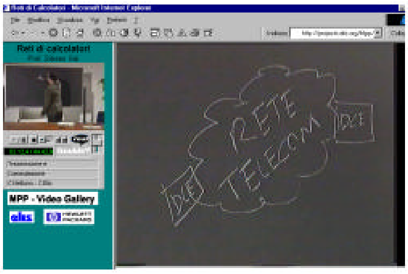
2.1.6 Other streaming systems

Last year in our site [10] we reviewed a dozen different systems. After some acquisitions by RealNetworks and by Microsoft, in September 1998 only the two major players are actually competing and they own nearly the whole market. Some other smaller companies are only filling specific niches (like video clips for e-mail) with good products.

2.2 The MPP format

On picture 1 you can have an idea of the overall appearance of the MPP format in an 800x600 monitor. The small sized (192x128) video stream runs at the upper left corner in a fixed frame. In the right frame a still picture (560x420) is shown. It is the reproduction of what the teacher is showing or has just shown on the video. The size is sufficient for a good reading, comparable to the one you get on TV with the original videotape. The still picture changes automatically every time the teacher uses a different picture (drawing, slide or whatever).

Picture 1



Picture 2

In picture 2 the teacher is drawing on the blackboard, and in the still frame you can see what he is going to show within a few seconds. The synchronisation of the video stream with the still image needs not to be perfect. There are transmission delays over the Internet that suggest triggering the still image a few seconds before it is shown. Their size varies from 10 to 40 Kbytes, therefore the time to show up completely (they are progressive JPEG, with medium-high compression that preserves details) goes from 5 to 20 seconds on a 28.8K connection. In the meanwhile the video stream does not stop because, through appropriate buffering, the connection can sustain both the transmission of the video and the image. The images are also stored in the local user's cache and can be retrieved immediately if the student rewinds the stream to re-watch part of the lesson. The video stream is not saved locally, unless the video server manager has enabled the "downloading while playing" feature and the user has got a Plus version (cheap but not free) of the RealPlayer.

2.2.1 The procedure

The final procedure for producing an MPP video is the result of an effort for achieving the fastest and easiest technique. The best solution for the encoding phase requires two computers, typically Pentium, at least 100 MHz. Table I summarises their characteristics. The names of the PC's are for easy reference during the description.

PC	Special hardware	Mission	Notes
ALFA	Osprey 100 video capture board or another supported by RealEncoder	Unattended encoding	Windows NT has better drivers for the Osprey than Windows 95. Double processors can be effectively used.
BETA	Any still picture capture board	Manual still picture capturing	Better if capable of capturing directly at 560x420 without need for resizing, and saving in progressive JPEG

Table I

While ALFA is encoding a videotape (the operator needs only to start and stop it), on BETA the operator first of all cuts, through a simple software tool in a few seconds, the unwanted head and tail (trailers) of the previously encoded file. Then, while fast forwarding the original cassette, the operator determines the relevant still images to be captured and proceeds to the still capturing, naming the JPEG files with a progressive number.
Meanwhile, by fast forwarding on the BETA the RealVideo file, the operator can take note, on a plain text file, of the exact timing, from the starting of the video, in hours, minutes and seconds, when the image should be triggered. After completing all images, a manual editing is needed to the timing file, for adding the complete URL of the related image, which is to be invoked during the streamed video. We wrote a simple program to accelerate this phase and to decide the amount of the anticipate triggering of the still image, as we explained in the previous paragraph. This amount varies according to the encoding characteristics: for a 20Kbps stream we

used 20 seconds, for a 50Kbps 10 seconds, for the CD version 2 seconds.
At the end of the selection of the still pictures, the timing file containing the URL's is merged with the streaming video file using a specific tool by RealNetworks to produce a synchronised stream.
We calculated that the manual phase of capturing images lasts on the average around one hour for one hour of videotape. It largely depends on the number of still frames to be captured. To this time you have to add a few minutes to set up the proper links on the web server. On the overall we have so far achieved a one to one time ratio for encoding a videotape into MPP format.
Choosing the proper still images is a personal evaluation, which is difficult to make automatic. Sometimes there are more than two frames per minute, because the teacher is showing rapidly different pictures. More typically there is one image every one or two minutes. An unattended scheduled encoding should be at a higher frequency than the maximum expected, theoretically doubling it. This would in any case overload the user with less important images while occupying more bandwidth.

2.2.2 The storage needs

It's very easy to calculate the storage needs for a video server, when it is based on streaming video. If the content is encoded at 20Kbps, it means that around 2.5 Kbytes are streamed every second, which means 9 Mbytes per hour. If the encoding is at 50Kbps, one hour needs 22.5 Mbytes, which becomes 45 Mbytes for a 100Kbps typical Intranet encoding. If we consider that in the worst case there is an average of one 30Kbytes image every minute, we have an extra need for the JPEG files of 1.8 Mbytes per hour. All the other HTML files for the support of the system are insignificantly small. We therefore conclude that in a 600 Mbytes CD-ROM we can deliver 30 hours of lessons at 40Kbps. We leave the CD remaining space for a copy of Netscape Navigator 4 and RealPlayer 5, both free of charge.
With the affordable price of hard disks, a video server with 1,000 hours of lessons at 20Kbps is affordable for any institution. One 11 Gbytes disk would be sufficient.

2.2.3 The encoding parameters

The exact tuning of the encoding parameters is a delicate task. Given the available bandwidth (e.g. 20 Kbps for connection with a standard modem), you have to decide the amount to be allocated for encoding the audio and the video. We used 8.5 Kbps for audio (4 KHz frequency response) and 11.5 Kbps for video, setting "Sharpest image": it means that if the person who is talking begins to gesture or move around quickly, the frame rate decreases as the motion speeds up and increases as the motion slows down, while the image sharpness remains nearly constant throughout. The resulting average frame per second rate was 4. We consider that a sharp image of a teacher speaking is better than a fuzzy one, even though in this latter case the movements may be more natural and smooth.
Allocating 8.5 Kbps for audio may even seem too much, compared to the video portion, provided that 6.5 or even 5 Kbps rates give fair results for voice only. Anyway, as the spoken part is a relevant one in delivering a lesson, we feel that we have to avoid any sound artefact produced by lower rates codecs, even if they are very light. The RealSystem G2 technology should allow the same good quality with lower bitrate.
Working with different material, such as videos regarding environment, experiments in electronics or physics, or nursing practice, requires a different approach. You cannot loose some moving details, therefore the size of the video must be greater. In these cases an Intranet or CD version is more suitable, because we found that at least 100Kbps are necessary for a reasonable quality.

2.2.4 The local test

In order to test both the infrastructure and the methodology, we run a benchmark with the Communication Manager students of ELIS Post-Secondary School in Rome. The course delivered was "Computer networks" by Silvano Gai, recorded on video by Consorzio NETTUNO [11], selecting 11 out of 40 lessons. We randomly chose 11 students from the 22 in the classroom and assigned each one a multimedia PC. The 11 computers (Pentium 100) were on a 10Mbit Ethernet network connected to a Hewlett-Packard E-30 server

(Pentium 133) running Windows NT with 64 Mb RAM and Real Basic Server 5.0. On the same network, at the same time, around 50 other users were working with other servers, so the load on the network was fairly high.
The other students were all together in an another classroom watching a TV set. Both groups had one hour for each 45 minutes lesson, so they could rewind and watch again a small part of it.
At the end of the hour, a 15 minutes multiple choice test was submitted to both groups, for individual answer. At the end of the 11 lessons two final tests were submitted. Some questions regarding their attitude and knowledge were added to the technical part.

2.2.5 The remote test

In order to assess the system from a wider audience and from a specific user point of view, we advertised it in the NETTUNO Consortium web pages, so that their students could test it and write their comments. As we also produced a full version of the 40 lessons course in a single 600 Mb CD (encoding at 40Kbps), we conducted a survey to understand whether this alternate delivery would be better than the on-line service. The questions were:

- Are you interested in buying a CD containing all the lessons (normally 40) of one NETTUNO course (like "Computer networks") in the MPP format?
- How much do you think you should spend for a CD containing a full course in streaming format?
- What kind of use do you plan for such a CD?

By registering the MPP site in the most popular search engines over the Internet, we also achieved hits from all over the world.

3. RESULTS AND DISCUSSION

3.1 The test results

The result of the comparison between the two groups of students, one watching the traditional videotape on TV and the other one using individually a PC for the streaming version in the MPP format, showed that there was no significant difference in their comprehension level.
In table II we compare the average points scored by the questions that were asked at the end of each of the eleven lessons. Points range from 1, meaning very little, to 5, meaning very much. The starting knowledge level of the two groups was even, so there wasn't any bias in the comparison.

Question	TV	MPP
I already knew the subject	2.2	2.1
The subject was interesting	3.6	3.2
The subject was clear	3.5	3.4
I understood	3.4	3.6
The teaching system is effective	3.6	3.4
I felt tired at the end	3.3	3.3

Table II

It is interesting to notice that the students in front of the PC did not feel more tired than their colleagues watching TV. The common tiredness caused by long exposure to a PC monitor was not noticed in a lesson lasting one hour. Even though a longer test should be accomplished before assuming that there isn't any additional difficulty with a PC instead of a TV, we have noticed that the way the students are behaving in front of a streaming video in the MPP format is a bit different from their behaviour with a multimedia hypertext, which requires much more user action. That's why we believe that, from a psychological point of view, the MPP format has an advantage to other PC-based educational material, retaining all the good of a TV

transmission and adding some more capabilities, like saving or printing the drawings or slides shown by the teacher, linking the video to questionnaires and other resources, implementing indexing capabilities.
The network performance values were good. Only in very few cases there were minor delays in delivering the streams. Some students complained that jumping back and forth in the stream caused always a few seconds rebuffering, which was annoying. The time lost in this rebuffering is, anyway, much lower than the time needed for a videotape to rewind or fast forward.

3.2 The survey responses

We had 4,500 visitors in six months on the MPP web site. As usual on the Internet, they were not only Italians. Actually the site is mostly written in English. Around one hundred (more that 2%, a fairly high rate for a web site) left their comments on the guest book: all of them were positive. A link with ad advice in Italian for the NETTUNO Consortium students, leads to a brief explanation of the project and invites them to fill in the survey. There were 52 responses from February to April. Some of the answers are resumed in table 3 and 4.

I WOULD PAY FOR THE CD	N. of answers
No answer	8
Less than 30 EURO	24
Between 30 and 50 EURO	15
More than 50 EURO	5
TOTAL	52

Table 3

TYPE OF USE	N. of answers
No answer	3
Diploma course	22
Personal study	22
Reference videolibrary	2
Other use	3
TOTAL	52

Table 4

The average sum the students would pay is a little less than 40 EURO. It seems that the solution of a CD is even preferred to an on-line system, because of connection costs. Anyway, with the recent reductions in telephone charges for connecting to an Italian Internet provider, the hourly cost has lowered to one EURO per hour in peak hours, and half at evening or holiday. So, a complete play via modem of the thirty hour “Computer networks” course, would cost between 15 and 30 EURO. Obviously, the CD gives the chance of watching many times the same lessons, without extra costs. The on line version is ideal for Intranet, and can be very useful for Internet connection, when frequent update of the content are required.

3.3 Copyright and authentication

The need of preserving copyright or allowing only registered users to access the video server is easily achieved through RealServer built in capabilities. They can also support a pay per view system.
In the MPEG-4 future standard there will also be capabilities for managing and protecting content owner rights.

3.4 Conclusion

This one year experience has proved that it is now possible for any institution to convert their existing video tape based courseware, or their recorded classroom lessons into a streaming format suitable for Internet or Intranet delivery, by using a cheap, fast and easy methodology. There are many applications for this technology: students can replay the lesson which they did not understand when they were present; absent students can recover the missed lesson; workers can improve their knowledge by following courses at any time, without any time constraint; very high level conferences can be available in a short time to everybody. The spreading of computers owned by educational institution as well as by the students themselves [12] makes this service an effective resource. Last, but not least, disabled people can access from home to resources which were previously available to students physically attending school or university [13].

3.5 Future enhancements

We are now implementing a mediateque, which is a streaming video server with searching capabilities and indexing of content, even inside the single video stream. The web site http://mediateca.elis.org contains more information and some freely accessible samples. The complete content is open only to the students of the "Linguaggi & Tecnologie Multimediali" course, which are experimenting the distance learning video based course on multimedia languages and technologies.

Research is in progress for automatic extraction of the still images, based on their persistence. Automatic indexing of the content is also being investigated through speech recognition algorithms.

4. ACKNOWLEDGEMENTS

The MPP projects is run by CONSEL - Consorzio ELIS per la formazione professionale superiore and CEDEL - Cooperativa Educativa ELIS. It is funded by the Italian Ministry of Labour and by Hewlett-Packard. We also thank RAI and Consorzio NETTUNO for their co-operation.

5. REFERENCES

1. M. Crudele and G Da Bormida: *Conversion of existing video teaching courses through Multimedia Production Process.* Proceedings of AICA 97, Sep 97.
2. http://projects.elis.org/mpp
3. E. Soloway: *No one is making money in educational software.* Communications of the ACM 41,2 (February 1998), 11-15
4. http://www.cselt.it/mpeg
5. R. Koenen: *MPEG-4 Overview - (Dublin Version), ISO/IEC JTC1/SC29/WG11 N2323.* http://drogo.cselt.stet.it/mpeg/standards/mpeg-4/mpeg-4.htm
6. http://www.apple.com/pr/library/1998/feb/11iso.html
7. V. Hardman, M. A. Sasse, I Kouvelas: *Successful Multi-party Audio Communication over the Internet.* Communications of the ACM, 41,5 (May 1998), 74-80.
8. http://www.real.com
9. http://www.microsoft.com/ntserver/netshow
10. http://projects.elis.org/mpp/players-comparison.htm
11. http://www.agora.stm.it/nettuno/
12. A. Deden: *Computers and systemic change in higher education.* Communications of the ACM 41,1 (January 1998), 58-63.
13. A. Scarfone, M. Crudele: *Internet project for disabled people*, in “Lo sviluppo tecnologico al servizio dei disabili”. SMAU, Milano 1998