Use of Audio and Video on the Internet

Richard Muirden

<richard@rmit.EDU.AU>

Abstract

Audio and video are finding increased use on the Internet as hardware and software become more widely available. However the Internet has its limits, and the use of high bandwidth applications pushes these limits to the extreme. Therefore attention must be given to the issues involved with use of these applications on the Internet.

1 Introduction

As of the last Internet Census the Internet currently contains over 30 million users and this figure is growing every day. With the Internet linking people together on a global scale, the usefulness of real-time audio and video between parties on the Internet is obvious.

However in reality, the Internet, as it currently exists, can only take so much traffic before Quality of Service (QoS) becomes unacceptable and other users are affected.

In the context of this paper, real-time Audio is defined as transmission of voice or sound such that it is reliably picked up at the receiving station in an understandable form with as little dropout as possible. Typically this is done with 64Kb/s Pulse Code Modulation (PCM) encoding.

Video transmissions, at current levels, are sent at 128K using a variety of formats resulting in the range of 3 to 5 frames per second (fps) achieved with colour video, and around 8-10 fps for monochrome video.

Although the actual usage of Audio and Video tools on the Internet would appear to consume less bandwidth than World Wide Web (WWW) and File Transfer (FTP) traffic, it is certain that more and more users will wish to use the resources of the Internet to achieve real time Audio and Video conferencing. Unlike private networks, where high speed and dedicated links are more common, the Internet is a shared inter-network with thousands of users doing many different things at any given time.

This paper examines the issues facing users of the Internet today and examines the need for change for the users of tomorrow with focus on the two most widely used systems on the Internet, the Multicast Backbone (MBone) and the widely used application for the Macintosh, CU-SeeMe.

2 Current Bandwidth Considerations

Currently there are over 40 products[1] for conducting real-time Audio and/or Video over the Internet. Some conform to international protocol standards such as the draft Real Time Protocol (RTP)[1] protocol or the CCITT H.261 video standard.[2] Others, however, such as the popular CU-SeeMe application use their own protocols - 4-bit greyscale video at 160x120 pixels or double that.[3]

With the wide mix of data formats and protocols currently in use, integration of services over the Internet becomes a problem. Some efforts have been made to allow for interconnection of the various formats such as the addition of code into the CU-SeeMe reflector to send audio streams out to a *vat* session, and video to a *nv* session for reception via the MBone.

Because of the large number of applications currently in use, and their methods of data transmission (eg: UDP/IP, IP Multicasting) there is no easy way to determine the total levels of use on the Internet. However we are already seeing strains on international links from sheer volume of traffic. In the future video and audio transmissions will comprise an even larger percentage of the bandwidth being used.

If we assume a “standard” audio transmission is using 64Kb/s PCM encoding, and a video transmission takes approx. 128Kb/sec, an average “conferencing stream” will take up to 200Kb/sec of bandwidth.

Bearing this in mind, the developers of the MBone chose 500kb/s as a target capacity. This assumes a minimum standard of a T1(1.5Mbps/sec) connection between sites. In practice, however, many sites connect into the MBone with more limited bandwidth available. IP Multicast, the protocol used by the MBone, enables one-to-many and many-to-many transmissions by using an algorithm that ensures that a packet crosses any one link only once.

However other applications, such as CU-SeeMe, which use unicast connections, can produce

[1] “Desktop Videoconferencing Survey Page” by Leigh Anne Rettinger. This WWW page is located at “http://www2.ncsu.edu/eos/service/ece/project/succeed_info/dtvc_survey/”

situations where there is wastage of bandwidth. For example: three people on the same network all connect to a CU-SeeMe reflector, located on another network, talk to each other, producing 6 unicast streams of data. If Multicast were used, the bandwidth usage would be considerably less, as only one copy of any packet to each recipient would leave the reflector and pass through links on the way to the destination network.

3 Uses and Users

Videoconferencing on the Internet takes many forms. From one or two people to thousands listening to a conference or rock concert. Unfortunately, as in most areas of society, there is some abuse of the Internet in the area of audio and video use. At least some would say it is abuse.

Remembering that the Internet is a "controlled anarchy" and is not owned by anyone, who exactly is to say what is a proper or improper use of the facilities? Some would argue that people who pay the bills are the ones to decide, but then again should it be up to any one network to influence material passing around the Internet? It is sometimes difficult to define what improper use is. This issue is outside the scope of this paper to discuss, however given the high bandwidth consuming nature of videoconferencing it would tend to affect more people.

For example, in late February, 1995, a user in Canada was found transmitting an X rated movie via CU-SeeMe to various receivers scattered around the Internet. Quite apart from any breach of copyright or obscenity laws in countries that the material was viewed in, this would be considered by most people to be improper use of the Internet.

However, in other cases the line is not so clear cut. Who is to say what is right and wrong? Van Jacobson commented that the largest increase of users on the MBone was due to the Rolling Stones concert of late 1994, while there are always listeners tuned into the NASA Select coverage, World Radio Network, and so on.

As part of the research for this paper a network survey of users of the CU-SeeMe application was conducted, in early March 1995, by the author. A total of 33 responses were received which, although a small sample, represents more general trends with responses from every corner of the globe. 36% said they used CU-SeeMe at least a few times a week, while 18% used it every day. Only 9% reported using the system rarely. 21% recorded they used CU-SeeMe when they heard something good was on. 60% said they used it for less than an hour at a time ranging to 12% who used it for 4 hours or more. 51% of users also sent video on a regular basis.

So even though the majority of users use the system for short periods of time, there are enough users of CU-SeeMe at any one time to possibly cause a network problem. If we extend the figures from the survey to generalise about use of other videoconferencing systems on the Internet there is some cause for concern about congestion of the Internet in the near future.

While 100% of respondents indicated that they would use a local reflector over an international one, it was surprising to note that 16% answered No to the following question: "When you use CU-SeeMe do you ever consider issues such as Network usage and loading, and bandwidth considerations?" Given that the question was loaded in favour of a Yes response, the high number of people who answered No would indicate that there is a significant amount of people who use the Internet for audio/video services who do not consider issues relating to bandwidth to be important. This suggests there is a class of user who will use as many resources as possible without due consideration of the results of their action. It is because of this that we need to formulate new strategies for coping with this kind of usage in the future.

If we couple this information with the fact that 94% of respondents used a non local reflector when using CU-SeeMe there is more cause for concern. In Internet terms, using a non local reflector crosses network boundaries, and in most cases users would be using the most popular reflector, based at Cornell University, which is where CU-SeeMe originates from. For some users this is right across the planet, and as such, overall Internet resources are being consumed, and consequently performance is affected.

4 Current Concerns

At the moment there are several areas in which audio and video usage lacks any clearly defined structure. This includes scheduling, resource allocation, standardisation and user education.

4.1 Scheduling

Currently there is no defined means of scheduling videoconferencing events. In the case of the MBone, some effort has been made in the form of a World Wide Web page known as the "MBone Global Agenda" where people wishing to broadcast may advertise their event by means of a form based page. Stephen Casner wrote on this very issue late in 1993:

> "There are no formal procedures for scheduling the use of the MBone, at least not yet. Management of the bandwidth is a cooperative procedure depending upon the good behaviour of the participants. [...] So far, there has not been

> any significant objection to the content of scheduled events. [...] We've been lucky so far that conflicts have not been a big problem. With a wider user community, we'd start getting conflicts between, for example, conferences for physicists and historians occurring at the same time. [...] So far, folks have been reluctant to establish more formal procedures (either to be the czar or to submit to one)."[2]

This applies even more so to use of other applications such as CU-SeeMe, the Internet Phone, and so on. As Casner stated we now have a much wider user community, and already conflicts are starting to occur.

When conflicts occur there is no procedure for dealing with them other than a "first come, first serve" attitude. However there have been instances where people have not announced transmissions and conflicts have occurred in real time, forcing protests from the MBone community. A recent example was the transmission, at around 600kb/sec, of a blank video image from Europe which interfered with another transmission taking place.

4.2 Resource Allocation

There is basically no method of resource allocation currently in use. RTP, the draft real-time transport protocol specification from the Audio/Video Working Group of the IETF "... does not address resource reservation and does not guarantee quality-of-service for real-time services."[1] Despite this, the IETF A/V Working Group is currently working on the Resource ReSerVation Protocol (RSVP) which, although still in draft, describes a protocol designed to "provide receiver-initiated setup of resource reservation for multicast or unicast data flows, with good scaling and robustness properties."[4] The basic idea is that when a receiver in a real time application (such as a videoconference) enters a conference, RSVP handles resource allocation throughout the route between the receiver and transmitter(s) effectively ensuring QoS for the entire data path. As useful as RSVP is, it does not address the issue of overload, nor is there any method to currently tell an application, or user that what they require from the network is going to cause significant problems bandwidth-wise.

4.3 Standardisation

Standardisation is another area where there is cause for concern. There is no agreed standard for either video or audio transmission. While there does appear to be a trend towards the CCITT H.261 protocol for video there are still other formats in use including the nv, cellb, JPEG, MPEG and CU-SeeMe format as well as other proprietary formats for various commercial products. Audio-wise there is the CCITT G.711 standard, along with PCM, Adaptive PCM (APCM) and DVI. Because there are so many standards in use, it is hard to integrate all the products on the Internet.

4.4 User Education

There is no real user education on the use of audio and video tools in the Internet environment apart from purely "how to" guides. Very little has been done to give users an understanding of issues concerning bandwidth and resource utilisation. Although some packages (ie: CU-SeeMe) include a note briefly warning users about using high bandwidth settings, these messages are buried deep in the documentation, and once the user enters the program, there is no "reminder" notice.

The MBone tools fare just as badly. Although the tools are still officially in beta test, more and more users are coming to the MBone with the impression of final release software. Provided documentation is minimal apart from standard format manual pages which, like those that accompany most UNIX tools, leave a lot to be desired in terms of being understandable by the non-expert.

Obviously application developers, protocol designers or Internet engineers should not be lumbered with responsibility for these things, but as the Cu-SeeMe survey has suggested, attention needs to be given to all of the factors mentioned above.

5 Do we need Video?

A question which has not received much attention is that of the viability of video streams over the Internet. Of course having a video picture is wonderful, but as a recent informal survey of the members of the Remote Conferencing Working Group found, most would prefer more reliable audio to a video image.

Considering that current technology allows a colour video image of 3-5 frames per second (fps) over the Internet as an average with 10 fps possible using greyscale versus "normal" television/film rates of around 30 frames per second, the question "How effective is this?" needs to be asked.

Of course for stable images 3 fps is fine. However for constantly changing images such as a concert, or a scene with many changes taking place, the frame rate just isn't enough and the viewer ends up with either a poor quality picture with portions of frames mixed in together, or a strobe like effect which misses any movement completely.

[2] Email message to the *rem-conf* mailing list, Dec 10, 1993.

As technology evolves, and bandwidth is not an issue, higher speed video will be standard, and this will not be a problem

Consider the application of a widely distributed conference. Video is not vital, and in fact past experience has shown that sometimes video can be useless, especially when an attempt has been made to focus cameras on difficult-to-read overhead slides.

Even a personal conference does not usually require full face-to-face viewing. The emphasis is usually on shared data (via a network whiteboard application, or similar device) rather than face to face contact.

White ISDN services and protocols are more specifically set up for videophone and other conferencing applications, Applications on the Internet should be directed more to wide area conferencing with an emphasis, at least currently, on reliable audio and shared whiteboard applications rather than video.

6 The Future

As of February 1995 the MBone has around 20,000 users on 1500 networks in 30 countries, with the numbers doubling every 8 months. If we take the explosive increase of the use of the MBone as a measure of the general trend, it is clear that at some stage in the future we will have a problem. The growth rate of bandwidth on the Internet is also growing, but it will take more time to catch up with demand.

Complementing the increase in audio/video users on the Internet is the increase in multimedia capable workstations. PC's are fast becoming capable of easily handling real-time audio and video while the price of the required hardware is falling. In essence we are seeing more people taking advantage of the opportunities that are opened to them.

In order to cope with this enormous growth, the Internet itself needs to be upgraded by several orders of magnitude. This is happening, but is expensive in terms of money and time. The evolution of applications and the increase of users is much faster than the rate of increase in available bandwidth.

Van Jacobson notes the enormous capacity of optical fibre being more than enough for any projected bandwidth requirements, however his projections were based on the continental US alone. If we factor in considerations of international links to the equation, which currently run over fibre, satellite, ISDN or other forms of communication links, then the picture looks a little different. It is true that the introduction of ATM and, in the future, broadband ISDN will make bandwidth less of an issue, but there is still the nature of global network connections to consider. The old saying about the "weakest link" applies here: It is useless having super fast ISDN to everyone's homes, if the connection between countries is swamped by a relatively small percentage of connections.

For example, the current link between Australia and the rest of the world runs at 3 T1 connections, or 4.5 Mbps/sec, or 20 simultaneous audio & video connections if unicast connections were being used. Australia has a Internet population of over 500,000. Put in these terms, that's not much bandwidth to share with other traffic such as FTP, WWW, email and news.

With this kind of situation in mind efforts must be made to introduce a generally agreed upon framework for the future use of the Internet for real time audio and video.

As noted above there is no easy way to control the usage of the Internet. Protocols will not, and cannot provide resource management and applications should only tell the user if the resources are not available, not that they should not be using them. As Van Jacobson points out, use (or abuse) of the Internet comes from users, and is a "social problem and we have no solution for it."[6] However several things to reduce the extent of the problem are: User education, Standard network usage such as IP Multicasting, along with the use of standard protocols, Scheduling and, as a final resort, enabling mechanisms to block unwanted traffic.

6.1 Standards

One of the current problems with the Internet is the plethora of Standards being used in all areas from mail systems to audio file formats. With the goal to get all users on the Internet "speaking the same language" Audio/Video applications should follow agreed upon standards such as ISO/CCITT or Internet Standards. While the development of protocols such as RTP helps with the inter-network transport of data, it is up to applications to use standardised formats with the end result that different applications on different platforms can talk to each other, and thus the end user benefits from this.

In the area of video transmission, the CCITT H.261 protocol (which is based on MPEG encoding) is fast becoming the preferred standard for video exchange. It allows for video streams to be compressed, sent and uncompressed in real-time with today's processing abilities of workstations. Although H.261 was designed with ISDN in mind, it appears to be the best protocol in use by applications for quality video transmission via a packet based network such as the Internet.

For audio there are many standards floating about, from the CCITT G.711 standard of 64 Kbit/sec 8kHz 8-bit PCM (and variants) along with DVI and GSM formats. Each has its own strengths and weaknesses (eg: Linear Predictive Coder (LPC) audio is best for transmission over slower medium such as modem links). The most used standard is however PCM (and its variations) and it would seem this is the standard to be followed.

6.2 Scheduling

As noted already there is currently no form of scheduling of conferencing events on the Internet apart from the informal system described. Clashes are becoming more and more common. An ideal solution would have a global scheduling system, which applications could recognise.

An example of this might be to broadcast scheduling information via the Internet Group Management Protocol (IGMP) that the session manager *sd* tool could handle. An extension could be made to allow for scheduling of events, and the notification of already existing events in the case of a clash.

In addition, there could possibly be some form of mechanism which would then control outgoing bandwidth of applications when it was known that other sessions were under way, perhaps in unison with a priority system.

Such a priority system could enable "background" sessions (for example, NASA Select sessions, which currently are always "pulled back" when they conflict with other sessions) could be assigned a priority that is low, while an IETF conference could be given a high priority.

These suggestions apply mainly to the MBone as it currently exists. How any form of scheduling could be brought into applications such as CU-SeeMe is unclear. It would seem that user education along with sensible bandwidth settings (see below) appear to be the only viable options.

6.3 Resource Usage

Applications currently allow users to select the amount of bandwidth to use (and as such, proper user education needs to be done - see below) while the administrative ends of applications such as multicast routers and CU-SeeMe reflectors allow controls to be put on the bandwidth transmitted. However as more and more use is made of the Internet a change in overall structure is needed.

The MBone is a good start. IP Multicast, being the most efficient use of the Internet resources to transmit audio and video to large numbers of receivers, should be adapted by other applications such as CU-SeeMe.

While there is small (but growing) use of multicast between CU-SeeMe reflectors (mainly as part of integration with the MBone), the current system of unicast links should be replaced with a more structured setup, along the lines of the MBone itself, with multicast links to deal with the increasing amount of traffic.

One proposal that is already in effect in Australia is to link all CU-SeeMe reflectors together via multicast links. This means that any user can connect via CU-SeeMe unicast to their local reflector, and share with all the other reflectors in the country via their multicast links rather than creating unicast links to reflectors that are further away. Although no data has been gathered from this exercise, the benefits of such a system are clear.

Currently most CU-SeeMe users, as evidenced by the aforementioned survey, connect to a non-local reflector (most use the reflector at Cornell University). If even one reflector per country were multicast linked to the Cornell reflector it is believed significant savings in network resources would occur. Such a setup would also mean faster response times for users.

A concept that has been introduced by the European MICE Project (Multimedia Integrated Conferencing for Europe) is that of "minimum acceptable quality" whereby "users specify a minimum quality of service (in terms of bandwidth, ... and various other parameters), and if those parameters and those of any other pre-booked conference cannot be met, the booking is refused."[5] This means that the user will get at least what they wanted (or possibly better quality), but not any worse than specified. While this approach is best suited to ISDN based conferences, the idea can be extended to the Internet by keeping statistics on link performance, and denying service if current link levels are below acceptable for the request made..

It is important for applications designers to make sure that it is more difficult for users to accidentally cause bandwidth problems by making the controls either only accessible via an "expert mode" or perhaps out of reach of normal users all-together. Naive and/or inexperienced users can easily cause network floods without realising what they are doing. Setting the slider to "max" isn't always a good thing. This is as much a issue for applications programmers as it is for users.

6.4 User Education

Procedures such as discussed above all rely on correct use of applications. Referring to the CU-SeeMe user survey, 16% of respondents indicated

they paid no thought to use of network resources or bandwidth when they connected. As the user community grows larger with more and more people who are not so-called “Network Hackers” but are more the “plug and play” people this figure will increase. as growth will mainly comprise the less computer literate.

Even though users should not need to worry about the complex configurations of the Internet, or bandwidth or resource usage, they should be aware of the consequences of their actions.

More attention should be given to informing users that moving settings outside of defaults will affect the quality of the transmissions they are sending or receiving as well as . For example, pushing an application to send at 1 Mbit/sec will produce more video throughput, but will totally consume and overload a network with only a 512Kb link into the Internet.

Also users should be educated on the global effects of Internet usage - the idea that when they switch on they are sharing, and using, tens or maybe hundreds of different network connections as part of their use of the Internet.

While education is critical in all areas of use of the Internet, it is important for users of high bandwidth applications such as audio and video to be aware of the resources they consume.

6.5 Blocking Mechanisms

Even with user education there will still be situations where control of Internet resources will be required. Currently it is very difficult to block transmissions from hosts or networks because of the connectionless nature of UDP packets.

As a last resort network administrators should be able to have some method for blocking high bandwidth traffic.

Currently CU-SeeMe reflectors offer configuration options to prevent streams from certain sites, but there is no easy way to prevent multicast packets from transiting your router. As Van Jacobson says:

> “Since a multicast talker can send nothing unless there’s at least one receiver, ‘stopping bad guys’ is already part of the architecture - don’t listen and they go away.”[6]

Currently the MBone is seeing experimentation with protocols to automatically prune and graft sub trees, and multicast packets can be limited by their time-to-live (ttl) value. [7] There is, however, no easy way to currently deal with the situation where you cannot stop your users from entering a multicast session. Although wanting to disable traffic might be seen as an unpopular or draconian move by network administrators, consider the following situation: A user connects to a session that is transmitting high bandwidth signals into the network. The transmission is swamping the local network and causing problems for other users leading to complaints. The conferencing user is not contactable by the administration, or refuses to leave the session. In this case a method to block selected multicast traffic would be useful.

However, as stated above the use of such a device would be as a last resort to the aforementioned strategies for Internet Audio/Video use.

7 Conclusions

Video and Audio usage on the Internet will continue to grow in popularity. As use spreads outside of the “core” computing professionals more care will have to be taken in order to keep the Internet functioning effectively.

Strategies need to be implemented to cope with the increasing use, with the biggest emphasis on user education, while making more efficient use of resources on the Internet by use of IP Multicast instead of unicast links.

Providing some form of Internet based scheduling is greatly needed to avoid clashes as well as providing a means of monitoring overall use.

Getting applications to talk to each other with standard protocols and data formats will widen the reach of Internet based conferencing as well as allowing for more inter-operation between dissimilar networks and programs. Use of Internet standards such as RTP and RSVP will further increase the effectiveness and efficiency of Internet based videoconferencing.

References

[1] H. Schulzrinne, S. Casner, R. Frederick and V. Jacobson, “RTP: A Transport Protocol for Real-Time Applications” Internet Draft, March 13, 1995.

[2] CCITT, “Video Codec for Audiovisual Services at p x 64 kbit/s - Recommendation H.261” Geneva, 1990.

[3] P. Hein and H. Kourous, “The CU-SeeMe Cookbook v1.0” Frequently Asked Questions (FAQ) Document, June, 1994.

[4] B. Braden (Ed.), L. Zhang, D. Estrin, S. Herzog and S. Jamin, “Resource ReSerVation Protocol (RSVP) - Version 1 Functional Specification” Internet Draft, March 1995.

[5] M. Handley, P. Kirstein and M. A. Sasse, “MICE Project Description” Computer Networks and ISDN Systems Volume 26, 1993 pp 275-290

[6] V. Jacobson, “The MBone - Interactive Multimedia on the Internet” University of California at Berkeley Seminar, Feb. 17, 1995.

[7] M. Macedonia and D. Brutzman, “MBone Provides Audio and Video Across the Internet” IEEE Computer, April 1994, pp 30-36.

Author Information

Richard Muirden obtained his B.App.Sci (Computer Sc.) from the Royal Melbourne Institute of Technology in 1992. Since then he has worked as a Systems and Network Administrator for the RMIT Computer Centres.

Address: Richard Muirden
RMIT Computer Centres
P. O. Box 2476V
Melbourne 3001
Vic.
AUSTRALIA

Phone: +61 3 9660 3814
Fax: +61 3 9663 5652
E-Mail: richard@rmit.EDU.AU