The MIM Web Gateway to IP Multicast E-Meetings*

Roland Parviainen^a and Peter Parnes^b

^{ab}Division of Media Technology, Department of Computer Science and Electrical, Luleå University of Technology, 971 87 Luleå, Sweden

ABSTRACT

As video conferencing and e-meeting systems are used more and more on the Internet and in businesses it becomes increasingly important to be able to participate from any computer at any location. Often this is impossible, since these systems requires often special software that are not available everywhere or impossible to install for administrative reasons. Many locations also lack the necessary network infrastructure such as IP multicast. This paper presents a WWW gateway system that enables users to participate using only a standard web browser. The design and architecture of the system are described and performance tests that show the scalability of the system are also presented.

Keywords: Multimedia gateways, Video conferencing, Multicast, RTP, HTTP, WWW

1. INTRODUCTION

Video conferencing and e-meetings systems are today commonplace on the Internet in several environments, especially in businesses that want to save money on travel costs or time spent on going to physical meetings. These systems are used for meetings, presentations and to continuously support for collaboration in a group setting, providing among other features text messaging, audio and video, and shared collaboration tools. Unfortunately it is impossible to participate in a video conference or an e-meeting session in many locations and occasions for various reasons. For example, many public access Internet terminals at e.g. web cafes, hotels or airport terminals offer only web access through a web browser and do not allow users to install custom software. Audio and video capture devices might also be missing. In other locations the access network might not support the necessary protocols such as IP Multicast.

One way to enable participation in e-meetings in a wide variety of different locations is by connecting e-meeting sessions to the World Wide Web. To be able to support as many systems as possible, the interface must be kept simple and limited to basic web standards such as HTTP/1.0¹ and HTML 3.2² or 4.01.³ While this does not enable a full two-way participation as audio and video input are limited it still creates a great benefit to mobile users. This paper presents the architecture and evaluation of a web interface gateway system to e-meetings that enables users to participate in an e-meeting session from any place through a standard web browser. The system is implemented using a commercial e-meeting tool, Marratech Pro.

2. BACKGROUND

At Luleå University of Technology collaborative workspaces are used daily in a wide variety of situations. Example uses include allowing students to view classroom lectures from home, enabling members of discussion groups to interact from a distance, and providing members of projects and research divisions with increased presence of each other throughout their work day. The last mentioned case is referred to as the “e-corridor”.

In the e-corridor members of the session can be both active or passive and the possible uses of the workspace range from giving or listening in a formal presentation to passive monitoring of video. Communicating with colleagues through the e-corridor often replaces other communication media such as e-mail, phone calls or personal visits. Users with broadband Internet access at home can choose to be part of the e-corridor, join meetings and follow presentations.

Copyright 2004 Society of Photo-Optical Instrumentation Engineers. This paper will be published in the proceedings of the SPIE Conference on Multimedia Computing and Networking 2004 and is made available as an electronic reprint with permission of SPIE. One print or electronic copy may be made for personal use only. Systematic or multiple reproduction, distribution or multiple locations via electronic or other means, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper are prohibited.

Further author information: (Send correspondence to R.P.)

R.P.: E-mail: Roland.Parviainen@sm.luth.se, Telephone: +46 70 5317170

P.P.: E-mail: Peter.Parnes@sm.luth.se, Telephone: +46 920 72421

Figure 1. Screen-shot of the different tools of Marratech Pro

2.1. Marratech Pro

The software used for the e-corridor is Marratech Pro, a commercial application by Marratech AB[†], which is based on earlier research done by the Media Technology division.⁴ Marratech Pro provides the users with the possibility to send and receive audio and video to and from other participants. In addition to audio, video and text messaging there is a shared whiteboard and a shared web browser. The chat, shared whiteboard and shared web browser tools all use proprietary protocols while the audio and video tools use standard protocols such as RTP as far as possible. It is designed to use either IP multicast or unicast. In the latter case traffic is tunneled through a media gateway called the “E-Meeting Portal”. The video streams from participants in an e-meeting is presented in the “participants” window, which gives a thumbnail overview of all the video streams currently received from the group, while a “focus” window displays the video obtained from a single group member with a higher resolution and frame rate. All audio streams that are not manually muted by the user are mixed and played synchronized to the video streams. The shared whiteboard is mainly used to make quick drawings, sharing applications by incorporating screen shots and sharing Microsoft PowerPoint slides during presentations.

While the e-meeting system provides many benefits in the daily work, it is often impossible get these benefits when out traveling or visiting somewhere and the application is not already installed or cannot easily be installed. The Marratech Pro system includes support for tunneling the traffic through firewalls and NAT network, but still requires the user to be allowed to install custom software and have the network capacity to receive the full session. Special versions that do not exist today would also be necessary on devices such as PDAs. A web interface is one solution to the problem of making e-meetings accessible from anywhere where we have a computer connected to the Internet. A web interface would also be useful as a light weight tool that can be used to get a quick overview of what’s currently happening in an e-meeting.

3. RELATED WORK

The WebSmile⁵ system is a gateway system for multicast video. It only supports the Motion JPEG codec and does not support any audio codecs. Apart from not having any Java video player applet, the system described in this paper uses the same ideas for video display. Mediascape⁶ provides a similar system for video as WebSmile. In Mediascape users can also leave messages to other users through a system called “Post-it” and it is possible to start a direct video call with other users. NYNEX Portholes^{7,8} is a web based group awareness tool which provides a web page with video snapshots and activity meters, which is also used in the MIM web interface.

4. E-MEETINGS AND THE WWW

Although HTTP and HTML are commonly used for strictly asynchronous media, several extensions and plug-ins for synchronous media such as video and audio exist, for example streaming media players and Java applets. In this paper we

[†]<http://www.marratech.com>

Figure 2. Architecture

will try to keep to the original features supported by HTML and HTTP as a way to get the widest possible support from the variety of web browser installations available.

A web interface or gateway to an e-meeting session can be implemented or deployed in two different modes:

Personal mode In the personal mode, the web interface is incorporated into the same application that the user uses normally, running on for example a desktop computer in the office. Each user accesses his/her own application to view the web interface directly or through a proxy at the edge of the local network if necessary.

Public mode The web interface can also run at a central location. All users access the same application, often making user identification necessary.

The MIM web interface described in this paper can support both of these modes. It is also possible to divide possible uses of the web interface into different levels depending on the amount of participation:

1. Passively viewing an e-meeting now and then to get a sense of group awareness
2. Participate by chat and the shared web browser and whiteboard tools
3. Listen to audio and see live video
4. Send audio and video, see full frame-rate video and the ability of to edit whiteboard pages

In this paper support for levels 1-3 are implemented requiring only HTML and HTTP support and an MP3 player. Level 4 requires more advanced browser plug-ins.

In the following sections we describe each media of Marratech Pro and possible solutions in more detail.

4.1. Chat

Chat can create problems depending on if the gateway is running in a personal or public mode. In the personal mode, public and private chat messages pose no problem since there is only one user of the gateway and the correspondent always knows who the user is. In the public mode, more than one user can send chat messages from the gateway. If user authentication is done by the gateway the user name can be added to all out going chat messages, making it possible to identify the sending user. This is sufficient for public chat messages, but there are still problems with private chat messages as it is impossible for the gateway to differentiate private messages from the e-meeting session to different users of the gateway if it is part of the session as one unique user. To solve this problem the gateway has to be able to join the session as different users, or the e-meeting software has to be modified to support this. Another low-weight solution is to create a convention where an e-meeting user adds the name of the recipient in private chat messages[‡].

[‡]For example, “[TO Roland] Hi!”

4.2. Shared whiteboard

The shared whiteboard tool is one of the most complex parts of Marratech Pro. A user can draw text, lines and other geometrical objects, import text and Microsoft PowerPoint and Word documents, etc. Although supporting the full functionality of the tool in a web browser would be impossible without using advanced web tools such as Java script or a Java applet, only displaying the whiteboard pages are much easier. The pages can be rendered as images which are then viewed in the browser. As a form of input, documents such as images or Microsoft Office documents can be uploaded to the gateway and distributed from there to the session.

4.3. Shared browser

When a user distributes a web page the complete page is downloaded and distributed to all Marratech Pro clients from the local cache using a reliable multicast protocol. The built in browser in Marratech Pro displays the original URL in the location field in the GUI, creating the appearance that the page was downloaded from that URL. There are two ways to get access to the shared pages from the gateway system: either the web browser gets access to the page from the local cache in the gateway or the browser uses the original URL of the shared page.

The complete page is sent over reliable multicast to all clients in Marratech Pro due to scalability: just sending the URL would create a HTTP GET request “explosion” as all clients would simultaneously request the same web page at the same time. This problem does not occur in the same way for a gateway system, as users have to manually follow a link to a web page.

4.4. Audio

Audio often is the most important part of an e-meeting or video conference, but it is also hard to support with only basic web standards as it is a synchronous media that requires continuous streaming. Several audio formats can be streamed over HTTP, such as Windows media or MPEG-1. MP3, or MPEG-1 Layer III,⁹ is probably the most well know and widely supported format. Most browsers need external plug-ins or programs to play these formats. A phone gateway such as the SystemBase Dialgate-2010 or the Cisco VG248 Analog Phone Gateway products could also be used to provide audio support.

4.5. Video

Similar to audio video is a synchronous and continuous media, but for video single frames of a video stream can easily be incorporated in a web page. This requires that the gateway decodes the video stream and re-encodes frames in an image format that browsers support. An illusion of a video stream can be achieved by continuously refreshing the image regularly, either by reloading the HTML page that includes the image or by updating the image with push-technologies. The experimental *multipart/x-mixed-replace* MIME-type¹⁰ can be used to push updates to for example images from the server to the client, but it is not implemented by all browsers⁸.

To support full frame-rate video streaming additional tools is needed. As in WebSmile,⁵ a custom Java applet can be used to create a video player in a web page. Plug-ins such as the Window Media, Real Video or Quicktime plug-ins all support streaming video as well. In both cases, the gateway has to re-encode the stream into the new format and stream it over HTTP, unless original format is supported by the player.

Fortunately, we might not need to support the full frame rate of the live video streams: in¹¹ the required frame rate for video conferencing is reported as being no more than at least 5 fps. If we are only interested in providing awareness,⁷ states that a frame rate as low as one snapshot every five minutes can provide group awareness in a work environment.

5. IMPLEMENTATION

Building on an existing application has several benefits such as:

- Easy access to already decoded and mixed audio and decoded video frames. All audio and video codecs supported by Marratech Pro are automatically supported by the web interface.
- Easy access to whiteboard pages as low level graphics objects that makes it easy to encode them to various image formats.
- The user only has to run one application.

Figure 3. Simple web interface

Figure 4. Simple web interface

Figure 5. Advanced web interface

- All input from the web interface to a session always comes from the same origin as if the user would use the normal tool.

On the other hand, this also means the gateway always requires the full user interface of Marratech Pro and can be hard to get access to from the Internet if the internal network is separated with firewalls and proxies.

See figure 2 for an overview of the architecture of the system. The MIMGlue component is also responsible for communication with the MIM¹² system, where the name of the system originally comes from. All information to and from the HTTP server are handled by this component. In later section when it is said the HTTP server retrieves information from Marratech Pro it is implied that it goes through the MIMGlue component. Both the HTTP server and the MIMGlue component are running in the same Java process as Marratech Pro.

5.1. HTTP and HTML

The HTTP server that is used is included in Marratech Pro and originally based on an HTTP proxy by Diego Cassinera written 1996, although not much is left of that code now. It only supports HTTP/1.0 and each new request is handled by a new thread. Basic HTTP authentication was added to the server for the web interface.

Images are encoded as either JPEG or PNG images; for video images JPEG offers higher compression than PNG, while for whiteboard pages and activity indicators PNG often offer better compression depending on the contents. The defaults are set in the Marratech Pro preferences. The audio is encoded to mono 16Kbit/s, 22KHz MP3 by using a JNI interface to the Lame¹³ MPEG audio encoder.

Almost no HTML code is built in to the software; everything is handled by a template and tag system. HTML template files are stored on disk for the main starting page, different components such as user details, chat and whiteboard pages. Tags in the HTML are replaced with suitable HTML code before web pages are served to clients. This process can be recursive, i.e. HTML code inserted by a tag can contain new tags. Care has to be taken so the recursion is guaranteed to end. The template system means that the complete look and feel of the interface can be changed by just editing HTML code.

Two versions of the templates are currently used: one simple version which does not utilize HTML frames or automatic refresh of content and one advanced version that through frames creates a view that looks more like the normal mode of

¹²Unfortunately it remains unsupported in the most common browser today, Microsoft Internet Explorer.

¹³<http://lame.sourceforge.net/>

regular Marratech Pro. Different frames such as the video snapshot display and the chat are automatically reloaded to show new content by using a *<META>* tag in the HTML header. The *multipart/x-mixed-replace* push feature is not used by default due to the lack of support for it in Internet Explorer, but it can easily be added by a simple change to the templates. Figures 3 and 4 show a simple version of the interface, while figure 5 shows the advanced version.

5.2. Activity indicators

As seen in figure 3, an activity indicator is added to the individual user views to make it easier to get an overview of the recent activity of a user since we cannot rely on the video stream to provide the information when video updates are far apart. The algorithm for detecting activity in video is similar to the one used in NYNEX Portholes.¹³ Different activities such as sending audio or chat messages are also added to the indicator in different colors.

5.3. Different Marratech Pro media

Chat Chat messages are intercepted in the Marratech Pro application and forwarded to the HTTP server which stores all messages. When a request for the chat component is received, a HTML form is created and all messages are inserted. Chat messages to the session from the web interface are entered in another HTML form, and transmitted to the HTTP server with a HTTP GET request. The message is then forwarded to the Marratech Pro application which sends it to the session and adds it to the chat panel in the user interface. If the web interface is running in public mode and user authentication is enabled, the login name is prepended to all outgoing chat messages so remote users can identify the sender.

Shared whiteboard The shared whiteboard HTML component retrieves the list of available whiteboard pages from Marratech Pro. The individual pages can be retrieved by getting a reference to the Java graphics objects that are used to show the pages in the user interface from which images can easily be encoded in different sizes. Documents to be shared with a session are uploaded to temporary storage at the web interface, and the file name is sent to the Marratech Pro application which distributes the new pages in the normal fashion.

Shared browser As new shared web pages arrive from a session, the URLs are stored in the HTTP server and listed when the Shared web HTML component is requested. When URLs are sent from the web interface to the session, the URL is sent to the Marratech Pro application which downloads the web page to the local cache and then distributes it to the session participants through reliable multicast.

Audio Requests for a live audio stream are registered with the MP3 encoder, and the thread handling the request waits for data to arrive to a local queue. Decoded and mixed audio are intercepted just before it is sent to the audio device. The raw audio data is placed in a queue, where the MP3 encoder retrieves it from. Encoded MP3 data is added to all currently registered HTTP audio request threads' local queue. As data arrives on the queue, it is transmitted over the HTTP connection. If queue grows too large data is discarded. If no audio is being received, silence is sent out to stop players from emptying their local buffers or rebuffering when there are short breaks in the audio.

Video When a request for a video snapshot is received, the data buffer of the last decoded video frame is retrieved, and encoded to the chosen image format which is sent back over the HTTP connection. The image can be scaled to show thumbnails as well.

5.4. Caching

To increase the performance, all requests to the HTTP server can be cached. Different types of objects such as video snapshots, whiteboard images or HTML pages can have different cache limits. To prevent web browser and proxies from caching the always changing information HTML *<META>* tags together with the *Pragma: no-cache* HTTP/1.0 header are used to state that the pages should not be cached. Different media have different needs when it comes to caching. As video snapshot often changes the cache should be short while the list of whiteboard pages does not change very often and can therefore be cached for much longer, although this is contrary to the performance gains of caching in the HTTP server as shown in section 6.

5.5. Privacy

If user authentication or identification is required, the Basic Authentication Scheme in HTTP/1.0 is used. If stronger authentication or encryption is needed, HTTPS (HTTP over SSL) would be the obvious choice. The HTTP server that is currently used does unfortunately not support HTTPS as of today, but SSL tunneling software such as Stunnel^{||} can easily be used together with the web interface.

^{||}<http://www.stunnel.org>

Figure 6. Bandwidth used for video images with different cache limits.

Figure 7. Number of successful image downloads per seconds with different cache limits.

Figure 8. Bandwidth used for HTML pages with different cache limits.

Figure 9. Number of successful HTML page downloads per seconds with different cache limits.

Figure 10. Time to download a video image with different number of download threads.

Figure 11. Time to download one HTML page with different number of download threads

6. EVALUATION

The test system used to conduct the evaluation tests described in this section were two PC computers on a local 100 Mbit/s Ethernet network. The client computer was running Microsoft Windows XP Professional on an AMD Athlon XP 2100+ CPU** while the server computer was running Mandrake Linux 9.1 on an AMD Duron 800MHz CPU. The server was also connected to the Internet through a 10Mbit connection and the Marratech Pro instance running the web interface was part of an e-meeting session on this network interface with a total of 14 users, where 10 were sending video. There were 5 shared web pages and 18 shared whiteboard pages.

The client software was a custom Java application designed explicitly for these tests. Both the client application and the MIM Marratech Pro Web Interface used the Sun J2SE 1.4.1 Java runtime. The test application is given an URL and the number of threads to start at startup. Each thread continuously tries to connect to the web interface through HTTP/1.0, download the data and disconnect as fast as possible. Statistics are regularly forwarded to a manager thread.

6.1. Audio receivers

With regards to MP3 audio over HTTP receivers, we evaluate the minimum delay experienced by a user and the scalability of the system with regards to the maximum number of concurrent listeners.

The audio delay experienced by a user is mostly affected by the amount of buffering done by the MP3 player: for example, the default settings of the Winamp v2.81 MP3 player the end to end delay is around 30 seconds. Using the same player, with the minimum allowed settings for HTTP streaming buffer sizes the end to end audio delay was around 4 seconds. This includes the time for encoding to 8Khz 64 kbit/s PCM audio stream in a sender on the same local network, transmission over RTP, decoding to raw audio at 32 Khz, encoding to MP3 and finally download and playout over HTTP. The relatively large delay is caused by several parts: MP3 encoder delay, buffering before data is handed off to HTTP clients, network delay and buffering by the client. Minimization of the delay has not been a goal so far in this work and is left as future work.

Scalability was only tested with up to 500 concurrent receivers. The audio delay for a Winamp 2.81 client as described above was still 4s. CPU usage on the server computer roughly doubled from around 20% to 40%, while the bandwidth used was as expected close to 1Mbyte/s (500 receivers at 16 Kbit/s). This shows that the only limits on the scalability of the number of audio receivers are either the network bandwidth or the amount of concurrent open sockets that the Java virtual machine or operating system can maintain.

6.2. Web server performance

All tests were performed with different cache limits, i.e. the time objects in the cache are still considered valid. When objects are invalid, they are regenerated. Both PNG and JPEG image formats were tested, since they offer some tradeoffs: the PNG encoder is faster than the JPEG encoder but produces larger files in general. The PNG video images were approximately 210 Kbyte while the JPEG video images were approximately 10 Kbyte. Three dynamically generated web pages were used for testing: the Corridor page, which lists the users of a session and shows a chat text area and was 4674 bytes large, the Shared web page, which lists the currently shared web pages and was 4733 bytes and the Whiteboard page, which lists the available shared whiteboard pages and was 511 bytes large.

In figures 6 and 7 the effect of the cache for the download speed and the successful number of downloads for both PNG and JPEG images with different amount of download threads is shown. The increase in performance with a large cache limit is clearly visible. The large size of the PNG images results in higher bandwidth usage but smaller amount of successful downloads. The number of concurrent download threads has a noticeable effect on the download speed and rate of successful downloads, where more threads in general reduces the performance of the server. Figures 8 and 9 show the same for the three different dynamically generated web pages. Here the beneficial effect levels out after a limit of 1s. The same performance drop when the number of current download thread increases is also visible.

Figures 10 and 11 show the effect of the number of concurrent download threads on the total time for downloading one image or one HTML page. In both cases, the download time grows linearly as the number of threads increases. The benefit of even a short cache limit is clearly visible for both images and HTML pages, but as above longer cache limits than 1s does not enhance the performance in the case of HTML pages.

**Running at 1733MHz

6.3. Availability

The default templates, as described above, have been tested with the most common browsers used today as well as text browser such as Lynx and Links. All features except the *multipart/x-mixed-replace* image push feature works in all tested browsers from Internet Explorer 3 and Netscape 3.04 and up, with the exception that images are not viewable in text browsers without an external helper program.

MP3 streaming over HTTP was tested with a variety of common players on different platforms, such as Windows Media player 7, XMMS, Winamp and iTunes. For some players the delay was substantial as the local buffer size could not be set to an appropriate value. While the delay would be too large for two way audio communication, it is often good enough for combined audio/chat communication.

7. FUTURE WORK

The next step is to examine and evaluate technologies to receive live video streams as well as live audio, by using formats such as Windows Media, MPEG-1, Real Video or Quicktime. As shown in section 6, the audio delay is currently quite large and work in minimizing it is needed. The possible use of advanced web technologies such as Javascript, Java applets and Macromedia flash will also be examined. The activity indicator displays was originally designed for small devices such as J2ME capable mobile phones, and should be redesigned as there is much more space available to use on a web page.

8. CONCLUSIONS

In this paper the architecture and development of a novel Internet e-meeting gateway have been presented. The system fulfill a need for users to be able to participate in multicast e-meetings while using limited terminals in places such as web cafe's and airport terminals and on computers where the user is not allowed to install custom software. It enables users to participate using only a standard web browser and supports a wide variety of different media.

We have discussed two different uses of the system: as a generic gateway into one e-meeting and as a personal gateway to a specific instance of an e-meeting client, and presented an evaluation on the scalability of the former alternative. Using a flexible template system, the appearance and the design of the web pages implementing the e-meeting can easily be customized, without requiring any change in the source code of the system. The evaluation shows that the system can scale to a large number of concurrent users if needed.

REFERENCES

1. T. Berners-Lee, R. Fielding, and H. Frystyk, "Hypertext transfer protocol – HTTP/1.0," 1945. IETF RFC1945.
2. D. Ragget, "HTML 3.2 Reference Specification," 1997. W3C Recommendation.
3. D. Ragget, A. Hors, and I. Jacobs, "HTML 4.01 Specification," 1999. W3C Recommendation.
4. P. Parnes, K. Synnes, and D. Schefström, "mstar: Enabling collaborative applications on the internet," *Internet Computing* **4**, 2000.
5. M. Johanson, "A RTP to HTTP video gateway," in *Proceedings of the tenth international conference on World Wide Web*, 2001.
6. N. Roussel, "Mediascape: A Web-based media space," *IEEE MultiMedia* **6**(2), 1999.
7. P. Dourish and S. Bly, "Portholes: supporting awareness in a distributed work group," in *Proceedings of CHI 92*,
8. A. Lee, A. Girgensohn, and K. Schlueter, "NYNEX Portholes: Initial User Reactions and Redesign Implications," in *GROUP'97, Proceedings of the International ACM SIGGROUP Conference on Supporting Group Work*, ACM Press, 1997.
9. M. Group, "ISO/IEC International Standard 11172; coding of moving pictures and associated audio for digital storage media up to about 1,5 mbit/s," 1993.
10. N. Borenstein and N. Freed, "MIME (Multipurpose Internet Mail Extensions) Part one: Mechanisms for specifying and describing the format of Internet message bodies," 1993. IETF RFC1521.
11. M. Chen, "Achieving effective floor control with a low-bandwidth gesture-sensitive videoconferencing system," in *Proceedings of ACM Multimedia 2002*,
12. R. Parviainen and P. Parnes, "Mobile instant messaging," in *Proceedings of the 10th International Conference on Telecommunications ICT'2003*,
13. A. Lee, K. Schlueter, and A. Girgensohn, "Sensing activity in video images," in *CHI 97 Extended Abstracts*, ACM Press, 1997.