The ePresence Interactive Webcasting and Archiving System: Technology Overview and Current Research Issues

Ron Baecker, Peter Wolf, and Kelly Rankin
Knowledge Media Design Institute, University of Toronto
rmb@kmdi.toronto.edu

Abstract: In contrast to video conferencing, webcasting supports scaleable Internet visual communications, yet it is typically viewed as an ephemeral one-way broadcast medium. We have developed a system design to support interactive webcasts that are accessible in real-time by remote viewers and retrospectively by archive viewers who can browse and search for what they want to see. ePresence is an open source system implemented using .NET technology that currently works with Windows and Linux servers and supports a wide variety of machines, operating systems, browsers, media engines, and bandwidths. We discuss current features of ePresence, its architecture and implementation, research intended to provide new capabilities, and the motivation for an open source release. Further details about the system's functionality, interface, and applications may be found in the companion paper "ePresence: An Open Source Interactive Webcasting and Archiving System for eLearning" by Rankin, Baecker, and Wolf.

The ePresence System

ePresence is an integrated hardware and software webcasting and event archiving system. To facilitate scaleable communications and knowledge sharing at a distance, we are engaged in research to make Internet visual communications:

- engaging, delivering rich media
- interactive
- accessible in real-time and via archives
- useful for knowledge building and sharing.

The result has been a viable and innovative webcasting infrastructure called ePresence (Baecker, 2002; Baecker, 2003; Baecker, et al., 2003). ePresence currently includes support for:

- video, audio, slide, and live computer desktop demos
- slide review
- moderated chat, private messages, and the submission of questions; and
- the automated creation of structured, navigable, searchable event archives.

ePresence also allows configurable live and archive interfaces through tailorable "skins".

Video, audio, slides, and computer desktops are automatically synchronized by the system. The "remote desktop" capability supports transmission of live 600X800 screen capture streams of live software demos and "Web tours" from the presenter's computer. Web links can also be sent by the speaker and synchronized with the video. Slide controls allow a remote viewer to review any slide already presented by the speaker. The chat system supports public chat, private messages, and questions to the speaker.

The archives interface allows retrospective navigation and browsing through a webcast using an outline of the logical structure of the talk and its slides and live demo sessions. Two levels of structure information are supported — chapters (sections of a talk) and slides. Slide titles are picked up automatically from Powerpoint in case it is used; chapter titles are input by the moderator during the talk and can if need be updated afterwards. Archive viewers can also navigate using a timeline. Additionally, we allow searching based on key words that appear in Powerpoint slides.

ePresence has been developed with full attention to valuable prior research described in Brotherton and Abowd (2004), Cadiz, et al. (2000), Hurst, et al. (2001), Isaacs, Mooris, and Rodriguez (1994), Jancke, Grudin, and Gupta (2000), Rowe, et al. (2001), Scott and Eisenstadt (1998), and Wactlar, et al. (1999). A more complete literature review appears in Baecker (2003). More detail about the system capabilities, interface, style of usage, and applications may be found in the companion paper Rankin, Baecker, and Wolf (2004).

System Architecture and Implementation

The system is implemented using .NET. The server software runs under Windows or Linux. Webcast can be viewed on client personal computers running the Linux, Windows 98/2000/2003/XP, and Mac 9.x or OS/X operating systems, and the Internet Explorer, Netscape Navigator, Mozilla, Opera and Safari 1.2 browsers, and using either Real Media or Windows Media live streaming. Archives may be produced in Real Media, Windows Media, and MPEG4 formats.

The architecture of our highly modular system may be portrayed as in Figures 1 and 2. The reader may find the following explanation of what all the modules do and how they fit together more understandable if she also refers to the discussion of features and interfaces that appears in the companion paper Rankin, Baecker, and Wolf (2004).

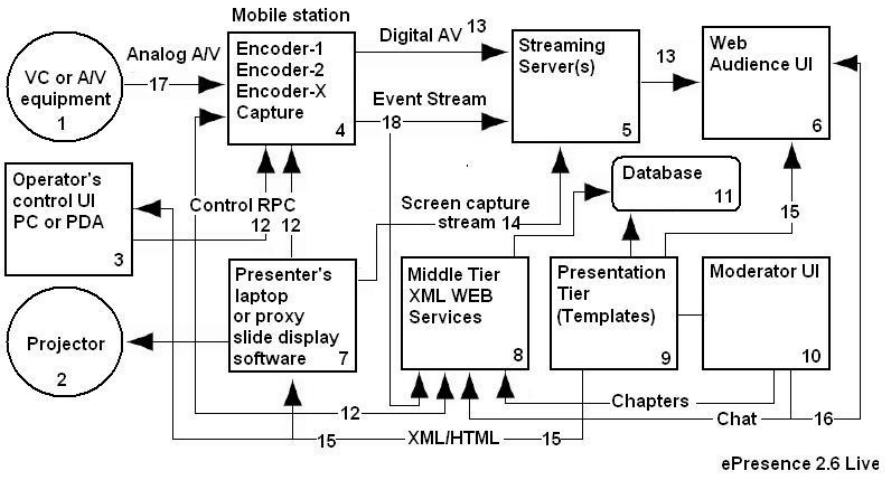

Figure 1: Current ePresence system architecture for live webcast

Interactive Webcasting (Live)

An ePresence live webcast is created by a speaker, an operator, and a moderator. These can be different individuals or the same person depending upon the scale of the event. The ePresence Mobile Station (4) includes several live media encoding and capturing software applications (e.g., Windows or Real Media) controlled by the operator (3) or speaker (7) via a single unified remote control interface. The remote control interface has been developed for different internet-connected devices (Laptops, tablet Pcs, and PDAs). The transport protocol used for remote control is an XML Remote Procedure Call mechanism called SOAP (Simple Object Accessing Protocol, http://www.w3.org/TR/soap/) (12). The operator can perform the following operations remotely: initiate live broadcast, start or stop archiving session, control slides transmission, submit URLs, and initiate multiple live software demo sessions. The speaker may give a talk to a local audience or remotely via a video telephone or videoconferencing (1). This allows us to webcast a meeting that is being held via videoconference. Web-based slide controlling and projecting (7) software allows having multiple distributed audiences listening and following the slide presentation in real time. The moderator interface (10) supports a local moderator who is watching the webcast, sending public announcements to a web audience, and submitting notes (chapter titles) to the archiving application. The moderator works as a communication "bridge" between the speaker and web audience transferring questions and comments on behalf of remote participants.

An ePresence webcast is typically viewed by both a local audience (2) and a live web audience (6). The web audience receives video and audio (13) of the speaker(s) from the streaming servers (5), a synchronized slide presentation stream (18) or a screen capture stream (14) from the presenter's computer, and web URLs (15). Remote viewers can also submit questions to the speaker (directly or via the moderator), have public or private text based dialogs (16), and review the slides that have been already presented. The presentation tier (9) is implemented as a set of web-based applications running on the .NET framework and several database engines (11). The server application has been recently ported from Windows to the Linux platform, and now supports Apache and MySQL software as a backend. The live interface (9) has been developed as a set of templates ("UI skins") that support

different layouts, media formats, video resolutions and other features. The operator can choose the most suitable template depending on the content of the talk. Adopters of the ePresence system can easily develop their own skins using XML, HTML, and several scripting technologies.

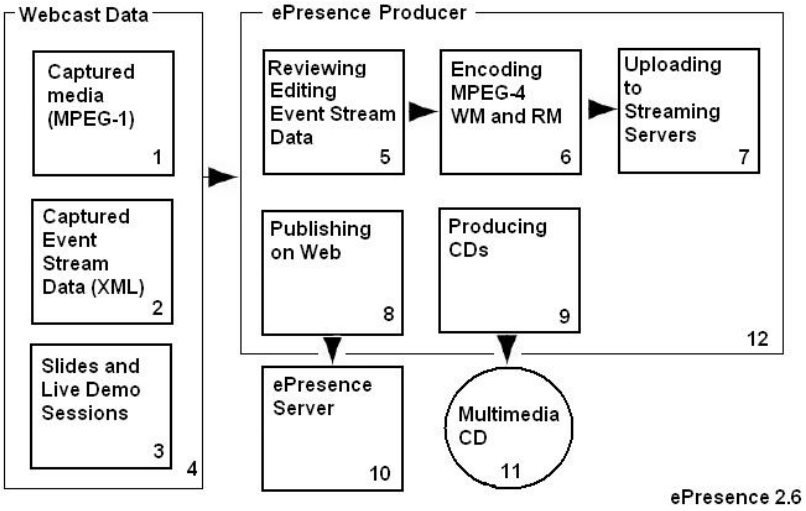

Figure 2: Producing an archive of an ePresence webcast

Archiving and Publishing a Webcast

The webcast data (4) such as video (1), slides (3) and event streams (2) is automatically captured during the live webcast. The events stream data includes time stamp information of slides and chapters submitted during the live webcast. Event streams can be updated (5) after the webcast using the ePresence Producer application (12). The operator can add additional keywords to enhance search, update slide synchronization data, edit chapter and slide titles, and replay the event with all synchronized materials before publishing the archive. The ePresence Producer software also allows encoding the captured video in different popular streaming formats (6), automatic uploading to a streaming server (7), automatic creation and publishing of web archives (8), and production of multimedia CDs (9). The software provides a selection of archive templates. The templates mechanism is based on XML/XSLT technology (http://www.w3.org/XML/, http://www.w3.org/TR/xslt). The published archive becomes automatically available on the ePresence website (10). It includes video player, slide frame, interactive timeline component, search tool, interactive table of contents, and threaded discussion board. Every archive exposes its keywords through the XML web services. This makes it easy to integrate the archives into different document repositories, "learning object" banks, and other searchable data storage systems.

Why .NET?

We considered several development platforms, environments and programming languages. The goal was to target the largest possible audience by developing a reliable application that can easily communicate with multiple 3rd party open source and proprietary video engines, streaming servers, media players and other components. Given the amount of computation performed by each element of the system, the solution had to be easily scalable. Communication among distributed elements required an open standard cross-platform protocol. Finally, applications needed to be written in a high level language available on different platforms to ensure portability.

The .NET framework meets all these requirements: it is based on open standards and available on all popular platforms. It is language independent. It allows one to quickly implement reliable and firewall-friendly communication layers for multiple distributed components using the XML-based SOAP communication protocol. For example, it proved to be an excellent vehicle for building the "Middle Tier XML Web Services" module (Module #8 in Figure 1) that lie at the heart of many of ePresence's capabilities. .NET also provides a good framework for 3rd party code reuse (e.g., COM on Windows, CORBA on Linux, JVM-CLR integration layer, and System Interop Services). It is available in both proprietary and several open source implementations (for Windows: http://microsoft.com/net/; for Unix-based platforms including the Macintosh: http://www.mono-project.com/).

Current and Planned Research

There remain many research challenges and opportunities, which may be characterized in terms of the degree of access by webcast participants, the richness and sense of presence provided by the media, the degree of interactivity, the characteristics and capabilities of the archives, and the quality and elegance of the system.

Providing More Participant Access

ePresence currently is only available to users of desktop or laptop computers. An implementation of a portable ePresence archive viewer on a Pocket PC has just been completed and will be tested by members of an undergraduate class this fall. We also plan work on supporting hand-held mobile access to allow local participants to participate in the chat, and to allow remote attendance via PDAs. More generally, a research project has begun with the goal of enhancing adaptation to user device and network capabilities by implementing a "Universal Multimedia Access" solution supporting the MPEG (4, 7, and 21) standards.

Enriching the Media

We seek to enhance the engagement and sense of presence experienced by remote participants, and to bridge the distance between local and remote participants. We are interested in how spatial (split screens) and temporal multiplexing (cuts and dissolves) enhance presence in webcasts. We are interested in how learning, attention, appeal, and stress vary with video quality.

Improving Interactivity

We have begun work towards reducing the delay between events and receipt of events. Upgrading to the latest Real Networks Helix and Microsoft Windows Media servers has enabled reduction in this delay from 25 to 30 second to 5 to 15 seconds. Our goal is a delay no longer than 5 seconds. We have begun the introduction of voice over IP using spatial audio (Kilgore, et al., 2003), and are experimenting with this a vehicle for participant discussion during an event, much like whispering to one's neighbor during a lecture. We will soon also allow questions to be spoken as well as typed. One research challenge here is to enable the smooth transition back and forth between the streaming mode used for the presentation and an audioconferencing mode for the Q&A and the discussion.

We intend to study the value of allowing remote viewing by groups as well as individuals. Gibbons, et al. (1977) showed that individuals viewing a video in a group learned more than those attending the live class who in turn learned more than those viewing a video by themselves. Smith, et al. (1999) have also reported similar conclusions using videoconferencing. We seek to learn if these results also apply to interactive webcasts.

Enhancing the Archives

We have carried out an initial study (Dufour, et al., 2004) of how viewers use structured, navigable, searchable archives, and plan more work of this kind. We have also begun research on further automating the production of structured searchable archives by automatically recognizing key words in the audio track (work has begun) and by using natural language processing to find topics. The voice recognition research will be informed by a Wizard of Oz study designed to elicit user requirements for recognition accuracy.

We plan the addition of threaded discussions over the archives. We conjecture and intend to test if this environment will encourage and support the formation of a "community" of online participants. We are particularly interested in how the online discourse enhances viewer understanding, and on how this depends upon the use of public chat and private messaging, and the integration of real-time chat during an event with later discussions over archived video. One aspect of commenting on video is the concept of collaborative annotations over video (see, for example, Tiernan and Grudin, 2001).

We plan to integrate essential features of the Expresto Creator digital video authoring, editing, and production capability (Baecker, et al., 1996; Baecker and Smith, 2003) into ePresence, which will allow the easy addition of titles, special effects, and editing of archived productions.

Improving the System

The major cost of using ePresence is now the camera operators and audiovisual technicians to produce quality webcasts. We plan research aimed at automating these functions, leveraging the work of Machnicki and Rowe (2002), Rui, et al. (2003) and Kapralos, et al. (2003).

Finally, we have discovered, as did Scott and Eisenstadt (1988), that there is a great need for flexibility in eLearning webcasting systems. Canned products rarely meet the great variety of needs encountered in different situations. To meet this need, we plan work supporting other eLearning standards such as SCORM (see http://www.adlnet.org/). More importantly, we have created a Consortium to release ePresence "open source." We conclude with a brief discussion of our plans in this area.

The Open Source Release

As of the writing of this paper (15 Sept. 2004), plans are well underway for release of the system by late in December 2004 or January 2005 under a dual license open source framework prior to the conference

Our motivation in going open source is to enable the formation of a community of institutions and individuals who can collaborate on future research and development and who can tailor the system to their particular needs. In addition to capabilities implied by the research directions just cited, here are some examples of the kinds of additions and adaptations that we would anticipate from members of our community:

- Modifications to the streaming engine so that they can provide an MPEG4 stream from a Linux workstation
- Integration of ePresence with one or more open source learning content management systems
- Addition of capabilities for tracking usage of video segments by students and using this in assigning marks
- Addition of quizzes based on material in the archives, and guidance for viewing based on the results of quizzes
- Development of an interface that requires minimal attention of the speaker so that a moderator is not needed..

We would also anticipate that other researchers would begin projects that appear on our "Planned Research" List, allowing the research community the benefits of multiple approaches to the same problem.

Additional discussion of the open source release appears in the companion paper Rankin, Baecker, and Wolf (2004).

Acknowledgements

Guidance of ePresence developments have since the beginning been shared with Dr. Gale Moore. Maciek Kozlowski of Videotelephony Inc. has provided continuous inspiration and many technical insights that have informed these developments. Maciek has been one of our primary sources of testing feedback; we are also grateful in this regard to Lawrence Spero, Rob Soucop, Martin Hicks, and Alessandro Tomassi. Ben Mazzotta has anchored the AV staff for major ePresence productions. Eric Smith has provided valuable code and technical insights. ePresence research is under the supervsion of Profs. Ron Baecker, Kostas Plataniotis, and Gerald Penn of the University of Toronto, and Prof. Elaine Toms of Dalhousie University. We are thankful for research support to the Natural Sciences and Engineering Research Council of Canada (NSERC) and to the Bell University Laboratories.

We wish to acknowledge the support and feedback from the following early adopters of ePresence:

- Videotelephony Inc.
- the Educational Computing Division of the Faculty of Medicine of the University of Toronto
- the Computer Science Department of the University of Trento
- the Health Sciences Information and Media Service of Memorial University
- Caseware International, and
- Informedia s.r.l.

References

1. Baecker, R.M. (2002). Highly Interactive Webcasting with Structured Archives, Poster Presentation. *Proceedings ACM CSCW'02*. Conference Supplement, 175-6.
2. Baecker, R.M. (2003). A Principled Design for Scalable Internet Visual Communications with rich Media, Interactivity, and Structured Archives. *Proceedings CASCON 2003*, IBM Press, 83-96.
3. Baecker, R.M., Moore, G., & Zijdemans, A. (2003). Reinventing the Lecture: Webcasting made Interactive. *Proceedings HCI International 2003*, Lawrence Erlbaum Associates, Volume 1, 896-900.
4. Baecker, R.M., Rosenthal, A., Friedlander, N., Smith, E., & Cohen, A. (1996). A Multimedia System for Authoring Motion Pictures, *Proceedings ACM Multimedia 96*, 31-42.

5. Baecker, R.M. and Smith, E. (2003). Modularity and Hierarchical Structure in the Digital Video Lifecycle. *Proceedings Graphics Interface 2003*, Halifax NS, 217-224.
6. Brotherton, J.A. and Abowd, G.D. (2004). Lessons Learned from eClass: Assessing Automated Capture and Access in the Classroom, *ACM Transactions on Computer-Human Interaction* 11(2), June 2004, 121-155.
7. Cadiz, JJ., Balachandran, A., Sanocki, E., Gupta, A., Grudin, J., Jancke, G. (2000). Distance Learning Through Distributed Collaborative Video Viewing. *Proceedings CSCW 2000*, 135-144.
8. Dufour, C., Toms. E.G., Bartlett, J., Ferenbok, J., and Baecker, R.M. Exploring User Interaction with Digital Videos (2004). *Proceedings Graphics Interface 2004*, London, May 2004.
9. Gibbons, J.F., Kincheloe, W.R., Down, K.S. (1977). Tutored Videotape Instruction: A New Use of Electronics Media in Education. *Science* 195, 1139-1146.
10. Hurst, W., Maass, G., Muller, R., Ottmann, T. (2001). The "Authoring on the Fly" System for Automatic Presentation Recording. *Proceedings CHI'2001 Extended Abstracts*, 5-6.
11. Isaacs, E.A., Morris, T., Rodriguez, T.K. (1994). A Forum for Supporting Interactive Presentations to Distributed Audiences. *Proceedings CSCW'94*, 405-416.
12. Jancke, G., Grudin, J., Gupta, A. (2000). Presenting to Local and Remote Audiences: Design and Use of the TELEP System. *Proceedings CHI'2000*, 384-391.
13. Kapralos, B., Barth, A., Ma, J., & Jenkin, M. (2003b) A System for Synchronous Distance Learning, *Proc. VI'2003*, Halifax, NS.
14. Kilgore, R., Chignell, M. and Smith, P. (2003). Spatialized Audioconferencing: What are the Benefits? *Proceedings CASCON 2003*, IBM Press, 135-144.
15. Machnicki, E. and Rowe, L.A. (2002). Virtual Director: Automating a Webcast. Proc. SPIE Multimedia Computing and Networking, January 2002. Mukhopadhyay, S., and Smith, B. (1999). Passive Capture and Structuring of Lectures. *Proceedings ACM Multimedia 99*, 477-487.
16. Rankin, K., Baecker, R., and Wolf, P. (2004, this volume). ePresence: An Open Source Interactive Webcasting and Archiving System for eLearning. *Proceedings ELearn 2004*, Washington D.C.
17. Rowe, L.A., Harley, D., Pletcher, P., Lawrence, S. (2001). BIBS: A Lecture Webcasting System. BMRC Tech. Report June 2001, Univ. California, Berkeley.
18. Rui, Y., Gupta, A., & Grudin, J. (2003). Videography for Telepresentations. *Proceedings CHI'2003*, 457-464.
19. Scott, P. J., & Eisenstadt, M. (1998) Exploring Telepresence on the Internet - the KMi Stadium Webcast Experience. In M. Eisenstadt and T. Vincent (Eds.). The Knowledge Web. Kogan Page.
20. Smith, R.B., Sipusic, M.J., and Pannoni, R.L. (1999). Experiments Comparing Face-to-face with Virtual Collaborative Learning. *Proceedings CSCL 1999*, Palo Alto CA., 558-566.
21. Tiernan, S.L. and Grudin, J. (2001). Fostering Engagement in Asynchronous Learning through Collaborative Multimedia Annotation. *Proceedings Interact 2001*.
22. Wactlar, H.D., Christel, M.G., Gong, Y., Hauptmann, A.G. (1999). Lessons Learned from Building a Terabyte Digital Video Library. *IEEE Computer*, Feb. 1999, 66-73.