Use Scenarios for Shared Editing in Scientific Collaborations

Marcia Perry
Lawrence Berkeley National Laboratory
One Cyclotron Road
Berkeley, CA 94720 USA
+1 510 486 6786
MPerry@lbl.gov

ABSTRACT
In this position paper, we describe our work in integrating shared editing with messaging systems for scientific collaborations within a widely distributed environment. We present scenarios and describe the requirements for collaborative editing integrated with group discussion and synchronous messaging. Our interest is in exploring research and development of shared editing systems to further our development of collaboration tools. We hope to inspire discussion of technology solutions for an integrated approach to synchronous and asynchronous communication and document sharing.

Keywords
Collaborative editing, file sharing, synchronous messaging, asynchronous communication

INTRODUCTION
Scientific collaborations are established to accomplish a wide variety of tasks within a wide range of organizational formality. These collaborations can be highly structured (e.g., analysis within high energy physics experiments) or informal (e.g., co-authoring a paper). Within these domains collaborations generally progress through various communication modes and durations. Often it is appropriate to exchange email messages but at other times it may be necessary to hold a scheduled videoconference. Some collaborations require maintaining continuous connections for on-going discussions and others are satisfied by exchanging ad hoc 'instant messages.' Sharing documents is often essential and the means by which this is accomplished can also vary considerably. Sometimes it is sufficient to merely view another person's publication and have a discussion via telephone or a messaging system. Other times it is necessary for several people to edit a document, independently or simultaneously. It may be sufficient to forward a file from one person to the next or to follow a paradigm in which a file is checked out from a repository, edited, and then checked in. Certain situations call for annotating or editing a document viewed by several users and propagating the changes to the remote viewers in real time.

Distributed computing enables large-scale scientific research by supporting interaction between geographically dispersed individuals, instruments, and resources. Grid environments are being developed to manage and provide access to a wide variety of computing resources such as high-speed networks, computer clusters, and mass data storage systems. An example is the DOE Science Grid[6], aimed at providing scientists access to heterogeneous and dynamic resources to process huge data sets and solve complex problems. The emergence of computational grid environments presents new opportunities and challenges for collaboration. This paradigm makes remote job submissions and file transfers routine operations. Therefore, in addition to sharing data, discussing analysis results, and collaboratively developing codes and editing publications, scientists will likely track workflow, view and monitor job status, and require a secure environment.

In order to facilitate these collaborative efforts, we envision a need for human-to-human communication mechanisms that are more synchronous and immediate than email but less intrusive and complex than videoconferencing. This has motivated our work on a messaging tool to enable collaboration among developers and scientists within and outside of the DOE Science Grid. This tool (LBNLSecureMessaging) supports text-based messaging on a group or one-to-one basis, and it provides presence and awareness information so that users can locate each other and determine availability. Conversations are mapped to venues, which may be public or private and permanent or temporary, where permanent venues can serve as rendezvous locations for particular interest groups.

To accommodate varying modes of collaboration, we are working to integrate file sharing and shared editing mechanisms into the messaging tool. We want to leverage existing systems and frameworks and incorporate this integrated communication system into a broader environment that we are developing--the Pervasive Collaborative Computing Environment (PCCE)[1]. We are concerned with integrating existing components into the PCCE. The next sections will present our usage scenarios and requirements for file sharing within scientific collaborations from the viewpoint of users who are in need of such systems.

USAGE SCENARIOS

Our use scenarios for scientific collaborations on the Grid are not very different from those in other CSCW environments (e.g., collaborative learning, group website creation, and corporate workgroups). However, they progress through various modes of document sharing associated with different means of communication, ranging from a single creator, multiple viewer model to group editing in real time.

For example, at any time and from various platforms, users may view a file containing status information from jobs executing on remote hosts or a graph output by a workflow manager. The readers may discuss the documents but do not edit them. This simple model is typically satisfied by a web-based approach in which files are stored on a web server viewed from a web browser via HTTP.

Other use scenarios involve independent or sequential group editing with optional discussion (e.g., global accelerator operators are writing a manual for the control room or researchers are co-authoring an agenda for a retreat or a paper for publication). The collaborators can work separately at any time or together at a virtual meeting. Traditional solutions follow the "retrieve from a central repository, edit, check back in" model with optional versioning and file locking, such as CVS-like systems or WebDAV[9].

A scenario depicting the need for annotating a file in real time or providing other visual cues during a discussion might be as follows: Ann and Bart are looking at a document containing a graph of analysis results. Using an instant messaging tool:

> Ann says: Don't you think the data point at (x,y) is curious?
> Bart says: What data point?
> Ann says: The point on page 3, graph on the bottom
> Bart says: This one?
> (and moves the cursor or makes a mark to indicate where he is looking)
> Ann says: Yes.
> Bart says: OK, looks weird. Think it's a bug in the program?
> Ann says: Maybe... Let's show this to Charlie.
> Ann invites Charlie, the programmer, to the discussion, and Charlie shares the same document.

REQUIREMENTS

Our system must support a variety of platforms (e.g., workstations, desktops, and laptops running Unix or Windows) and allow users to access documents in various formats (e.g., Word, Framemaker, or plaintext). Since users are located around the world and often travel, documents must be readily available at any time and the means of sharing them must be easy to use and ubiquitous. Collaboration should be just as possible from a computer borrowed at a conference or experiment site as from a home or laboratory desktop. Furthermore, we must support many modes of document sharing, including the "single publisher, many viewers" scheme, independent and asynchronous shared editing, real-time simultaneous annotation and editing.

Security requirements include user authentication, access control, and privacy. For integration into the PCCE, implementations must allow server and client authentication with X.509 certificates and encryption of data sent over SSL network connections. With either a client-server model or a peer-to-peer system with optional servers, collaboration groups should be able to run and configure their own servers at their chosen sites.

DISCUSSION TOPICS

Since flexibility is important, we view the PCCE as a composable system that allows integration of several tools such as messaging, file-sharing, and collaborative editing mechanisms. There has been much research and development on collaborative editing tools and frameworks (e.g., LotusNotes[4] and DOORS[8]). Collaborative website creation has been enabled by systems such as Zope[7] and CoWeb[5], and web annotations are supported by systems such as MicrosoftOffice2000[3]. Shared window systems and tools such as whiteboards allow real-time visual cueing. Other collaborative interaction environments have been developed to integrate shared editing with messaging. These include AOL InstantMessenger, the Collaborative Virtual Workspace, Groove, WebEx, and many others.

Role-based models such as Zope and BSCW[2] are attractive in their ability to provide access control. Our use scenarios and requirements argue for web-based solutions such as BSCW or web servers built with WebDAV support (e.g, Apache with mod_dav) that are accessed by clients such as DAVExplorer. We propose discussion that extends this survey to explore additional systems that integrate shared editing and real-time annotation with messaging.

REFERENCES

1. Agarwal, D., McParland, C., Perry, M, "Supporting Collaborative Computing and Interaction," *Proc. of the Grace Hopper Celebration of Women in Computing 2002 Conference*, October 9-12, 2002, Vancouver, Canada.

2. Appelt, W., "What Groupware Functionality do Users Really Use?," *Proc. of the 9th Euromicro Workshop on PDP 2001*, Mantua, February 7-9, 2001, IEEE Computer Society, Los Alamitos.

3. Cadiz, J., Gupta, A., Grudin, J., "Using Web Annotations for Asynchronous Collaboration Around Documents", *Proc. of ACM Conference on Computer Supported Cooperative Work*, December 2-6, 2002, Philadelphia, PA, p. 309-318.

4. Falkner, M., "Using Lotus Notes As an Intranet," John Wiley & Sons, March, 1997.

5. Guzial, M, Rick, J., Kerimbaev, "Recognizing and Supporting Roles in CSCW", *Proc. of ACM Conference on Computer Supported Cooperative Work*, December 2-6, 2002, Philadelphia, PA, p. 261-267.

6. Johnston, W., "Vision for Collaboratories and the DOE Science Grid," Presentation at the DOE2000 Review, May 2000, Argonne, Illinois.

7. Lattier, A., Pelletier, M., "The Zope Book", New Riders Publishing, 1st edition, July, 2001.

8. Preguica, N., Martins, J.L., Domingos, H., Duarte, S., "Data Management for Asynchronous Groupware," *Proc. of ACM Conference on Computer Supported Cooperative Work*, December 2-6, 2002, Philadelphia, PA, p. 69-78.

9. webdav: "IETF WEBDAV Working Group", http://ftp.ics.uci.edu/pub/ietf/webdav.