Planning, Building, and Using A Distributed Digital Library

William J. Adams
Department of EE & CS
United States Military Academy
West Point, NY 10996
Voice: (914) 938-5575
FAX: (914) 938-5956
adams@exmail.usma.edu

Bernard J. Jansen
Department of EE & CS
United States Military Academy
West Point, NY 10996
Voice: (914) 938-3233
FAX: (914) 938-5956
jansen@exmail.usma.edu

Todd Smith
Department of EE & CS
United States Military Academy
West Point, NY 10996
Voice: (914) 938-2200
FAX: (914) 938-5956
smith@exmail.usma.edu

Abstract

There have been many reports on the construction of digital libraries. We find three shortcomings in many of these reports. (1) They fail to articulate the organizational need for a digital library. (2) They ignore the important operational details in establishing the library, including the end user classroom. (3) They fail to provide any quantitative data concerning the ability of the digital library to deliver multimedia files, a primary objective of digital libraries. In this paper, we discuss the U.S. Army's Army Training Digital Library as a necessary component of the Army Distance Learning Plan. As we are well on our way to implementing this plan, we present lessons learned from the project so far, specifically in the areas of classroom construction and results from experiments with delivering multimedia files.

Please Cite: Adams, W. J., Jansen, B. J., and Smith, T. L. 1999. Planning, building, and using a distributed digital library. Third International Conference on Concepts in Library and Information Science. Dubrovnik, Croatia.

See Other Publications

Keywords

Digital libraries; distance education; video delivery

Background

Schools, businesses, and governmental organizations are turning to Distance Learning to bolster enrollments, share expertise, extend the geographical extent of training programs, and broaden their customer base (Dance 1997, Etter 1995). Within an educational environment, Distance Learning (DL) provides a means to keep faculty employed and low enrollment courses viable through video conferencing and digital libraries (Fox 1996, Palounek 1996). The mechanisms required to deliver an instructor’s video taped lecture or printed material to students at various locations and times are fairly well documented and routine. The challenge lies in formulating a method to distribute interactive, multimedia educational resources on large scale, in a timely and cost-effective manner.

Research is continuing on the most effective utilization of networks for DL (Stanford 1997). Current DL programs are using a mixture of 3.5" floppy disks and CD-ROMs to distribute multimedia files to students (Harris et. al. 1994, Lollar 1995). The 1.44 MB size restrictions of a floppy disk make it impractical for anything but text files or application files to be distributed. Likewise, CDs present two challenges to instructional developers. First is the 650 MB capacity limit of the compact disk. With a typical AVI file averaging 12 MB per minute of video, this storage limitation is quickly reached. Second, because of the "write-once, read-many" nature of compact disks, changes to any file on the disk, no matter how slight, require a new master and all of its accompanying charges, effectively erasing most of the savings realized through the use of CDs in the first place.

To provide learning materials quickly and efficiently to students, many schools have turned to digital libraries. Some institutions like the University of Minnesota, utilize a type of monolithic storage arrangement to provide learning material to a continuing education student population that is scattered over several hundred miles of the northern United States (Duin 1996). Others, like Virginia Tech use digital libraries as a reference store for local students (Fox 1996). While the exact methods of storage and retrieval may differ, both examples allow students access to repositories of multimedia information through some type of network connection. These server-based edifices allow any student with network access, either through a local area network or dial-up access, to access or browse learning material of any type. Because of the immediacy of server access by instructors and developers, the material is the most current it can be, without the lag time of disk mastering or disk distribution. As such, digital libraries become a natural extension of any distance learning plan.

Problem

There are some issues with operating and maintaining digital libraries, however. These problems originate with the diverse nature of the user population. The variables in the retrieval equation are the users' connection method, connection speed, knowledge of search queries, and preference of file formats. The enormity of this problem becomes more evident when one looks at the global nature of the users of Army training material.

Digital libraries must be accessible if they are to be successful. Therefore, where the information is stored and how these storage sites are interconnected is of concern. Where do we place the servers and how do we connect them to each other? Of equal importance is that material in the digital library must be arranged and organized so that users do not spend an inordinate amount of time searching for the material they need.

Background

The United States Army has made a significant investment into distance learning capabilities. The goal is for soldiers to be able to access training material from any

location. In 1995, the U.S. Army's Training and Doctrine Command (TRADOC) embarked upon a plan that is intended on "leveraging technology to improve training" at the 21 Army schools (US Army 1997). TRADOC is similar to a University and the 21 TRADOC schools are similar to a Department within the University. A significant difference though is geographical dispersion.

The 21 TRADOC schools are in various locations across the United States. TRADOC is responsible for developing the training doctrine, the training material, training standards for both military tasks and doctrine. Previously, TRADOC distributed this material via hard copy. The organization then moved to disk distribution; however, this soon proved unworkable on a global, Army-wide scale. TRADOC began exploring other options for distribution.

The long-term intent of the Army Distance Learning Plan (ADPL) is to provide high-quality, standardized training material to soldiers around the world. By providing access to this material, soldiers can receive training or review material in any combination of locations. First, attendees of resident courses will access training material in the classroom through local area networks. Second, upon returning to their workplace the soldiers can review the material from their local education centers, armories, Reserve Centers, or their offices by way of the Internet. In the near future, soldiers will use the Internet from their home, using their personal computer.

There are many ways to quantify the efficiency and cost-effectiveness of this vision. The most immediate means are strictly monetary. The cost of connecting users is undeniably less than the amount currently spent on: mailing and updating course books for correspondence and nonresident courses; travel and per diem costs for resident training; and the lost productivity of students that have to travel. The greater, but harder to quantify, benefit is the guarantee of standardized, on-demand training anywhere in the world. This benefit is especially important for the military, with a large, mobile, and geographically diverse population.

The key component to this architecture is the backbone connectivity that will be used to transfer the training material between school and student, provided in this case via the Internet. The Internet does not provide instant access to the DL; but it does provide a global path. Each DL on the network would contain a cached copy of material from other DLs. If the cached copy were too old or too big to store, then the local DL would request an update from the source DL. Within the ADLP, the training material that is transported over the network is a combination of multimedia files. Video, audio, and text are the three largest components of learning resources. Once these materials are retrieved from their source, they are temporarily stored at the network access point's Digital Training Access Center (DTAC.) From the DTAC, users can view and replay the material at their convenience, for their use or for a class.

The planned training environment redefines the concept of the training site. The training site could be a school classroom, a learning lab at a military post anywhere in the world, or a National Guard Armory. Regardless of its location, the site is equipped with a set of

hardware that enables a specified set of functions. These functions include Internet access, World Wide Web browsing, multimedia capability, and local area network connectivity.

The distance learning retrieval process works as follows:

1. A soldier arrives for training. He accesses the local DTAC and requests training material.
2. The local DTAC searches its index of material that is currently stored there. If it finds what the user needs, the DTAC notifies the user that requested material is available and returns a list of Uniform Resource Locators (URLs) of the appropriate files. The process now jumps to step 5.
3. If the requested material is not on the local DTAC, the request is forwarded to the ATDL to find the course material he needs. This material supports correspondence courses, individual professional development, on-the-job training, or review of material learned at a previous resident course.
4. The ATDL searches its database to find course material related to the soldier's request. It returns a list of material in the form of URLs of the appropriate files. The DTAC displays the list of material with estimated download times, calculated from the current network load. The user can select any or all of the material, which is then downloaded to the local DTAC.
5. Once this download is complete, the DTAC notifies the user and prepares for delivery.

Delivery is dependent on both media and the user. Text, either in the form of HTML or Adobe Acrobat files, is downloaded to the user via a Web browser. Audio and video files represent more challenges however, although streaming products have simplified their delivery.

The Distant End

Many discussions of digital libraries and distance learning architecture focus almost exclusively on the network and the servers. This is understandable considering that this aspect of the architecture is many times the most exciting. However, the distance end, i.e., the classroom itself is to the user the most important component of the system. For it is here that the student and professor will interact. It is here that the system must perform. Figure 3 is a computer model of our existing classroom, modeled in OPNET, a military networking modeling and simulation application.

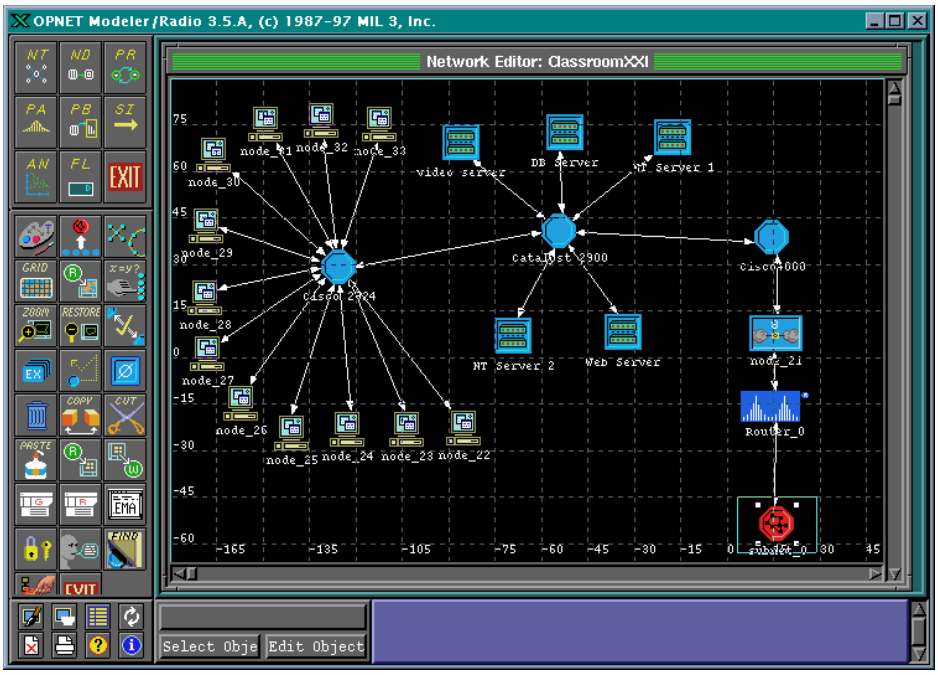

Figure 3 is a computer model of our existing classroom

Since the development of the classroom has received so little attention, many may be unaware of the expensive involved in setting up a classroom capable of supporting the wide range of digital library media and mediums. The media may include text, images, video, etc. The medium may be an Intranet in addition to a connection to the Web capable of supporting multiple and simultaneous accesses. Plus, there is all the normal support required for a professor in a classroom. Additionally, we have discovered that to take full advantage of the new technology and to emphasis small group exercises, some modifications to the standard lecture hall environment are desirable.

Summary and Conclusion

Digital libraries, whether distributed across the globe or across a small campus, provide a quick and cost effective means to distribute learning resources to students, employees, or soldiers. Choosing the best system requires an analysis of the organization, its location, its user population, and the type of material being provided to the users. There must be a realization that there is a cost in establishing a multimedia infrastructure. These costs include those of the classroom and the computing and networking infrastructure to support multimedia delivery.

References

Dance, M. (1997). *The Promise of Distance Learning*. http://Weber.u.washington.edu/~jamesher/mdance.htm.

Duin, A. H., and Nater, E. A.. (1996) *Designing and Managing Virtual learning Environments for Secondary, Post-Secondary, Graduate, and Continuing Education: A Land Grant Perspective*. Proceedings of the World Conference on Educational Multimedia and Hypermedia 1996. 202 - 207.

Etter, D.M.; Orsak, G.C.; Johnson, D.H. (1995) *A distance learning laboratory design experiment in undergraduate digital signal processing*. 1995 International Conference on Acoustics, Speech, and Signal Processing. Conference Proceedings. 2885-7.

Fox, Edward. (1996) *Digital Libraries, WWW, and Educational Technology: Lessons Learned*. Proceedings of the World Conference on Educational Multimedia and Hypermedia 1996. 246 - 251.

Harris, J.A.; Murden, C.; Webster, L.L. (1994) *The potential of interactive multimedia on CD-ROM to enhance laboratory work in physical science and engineering*. 1994 IEEE First International Conference on Multi-Media Engineering Education Proceedings. 296-301.

Lollar, R.B. (1995) *Distance learning for non-traditional students to study, near home, toward a UNC Charlotte BSET degree*. Proceedings IEEE Southeastcon '95 Visualize the Future. p. 366-7.

Palounek, A. P. (1996) *Distributed Computing Network for Science and Math Education in Rural New Mexico*. Proceedings of the World Conference on Educational Multimedia and Hypermedia 1996. 557-562

Stanford. (1997) *An On-Line Distance Learning System Using Digital Video and Multimedia Networking Technologies*. http://minas.stanford.edu/project/project.html. Accessed November 1997.

U.S. Army Training and Doctrine Command. (1997) *The Army Distance Learning Plan.* Available on-line from http://www-dcst.monroe.army.mil/adlp/adlp.htm. Accessed 1 December 1997.