LIBRARY EVOLUTION 3: MICROSOFT TO LINUX MIGRATION

INTRODUCTION

Our Context

Fletcher Library on ASU's West campus is visited by an average of 1,000 people each day who come to access the physical and virtual collections that support their research and studies. Maintaining a library has always been an expensive venture, but costs have escalated dramatically as libraries have moved from physical card catalog systems to "dumb terminals" in the 1990s to today's sophisticated computer workstations that connect users to the World Wide Web and a myriad of electronic collections. In 2005, we not only provide books, media, journals, and study space, but we must also purchase electronic subscriptions and the computing hardware and software that support all levels of research and learning. Historically, Fletcher Library has replaced its public computing hardware every three years. These three-year cycles have provided an opportunity to assess future needs for the next cycle replacement and, in 1999, as we faced budget revertments and rising costs, the need for change became evident.

The Problem

Since 1996, the Library has used Microsoft operating system and software applications. Our software licensing costs and new hardware costs were increasing. Our long-term technology support funding has never been stable since, like many departments within the University, we have relied on salary savings to fund technology-related purchases. As we filled vacant positions to meet the growing numbers and needs of Library users, flexible spending from salary savings was diminished to a level that made technology funding requirements unsustainable. An additional complication came in the form of budget revertments in FY 01/02.

Furthermore, Library services were limited by the Microsoft Windows operating system and applications. Microsoft is designed for home and business use, and source code cannot be altered in ways that support modifications useful to reference and instruction services. Instructions that would be most helpful for a user cannot be integrated. The Library Technology Support and Development Department (TS&D) was therefore not able to respond to staff requests to change menus, headers, dialogs, etc. of the Windows desktop and applications. These limitations and restrictions prevented us from providing students and other users with personalized or customized applications and interfaces that could enhance their learning experience.

System problems increased over time and required greater financial and staff resources. Troubleshooting consumed 18% of technology staff time as we dealt with problems which included unknown application and operating system crashes, system memory failures, hard drive corruption, and server patches, fixes, and anti-intrusion measures. Our users experienced inaccessibility of the floppy drive, default of our home page to porn sites, and disabled workstations - all resulting in disruption of their research.

Two things became apparent: Using the existing Microsoft architecture, the Library would not always have the resources to sustain its technology infrastructure and could not take full advantage of information resources and instructional opportunities. It was equally obvious that solutions would not be forthcoming from Microsoft. As we conducted research for alternatives, we identified a movement within the Fortune 500 companies that were making serious financial investments in open source-based technology solutions using Linux. Our research of university libraries revealed instances of only partial application of Linux in single servers, specialized workstations, or small labs. No models in production existed to show us the way.

To be objective, we investigated other Unix-like systems including BeOS, BSD, Mac OS X, and Sun Solaris operating system. After detailed review, Linux was identified as a viable and affordable alternative operating system.

The Solution

By 1999, TS&D staff had thoroughly researched the application of Linux technology through review of research and Internet articles and discussions with information technology peers in various countries including Australia, Brazil, Germany, Mexico, New Zealand, and the United Kingdom. After considerable exploration, TS&D proposed a drastic change in the Library's core computing architecture: to migrate our public and classroom workstations and servers from Microsoft Windows to Linux. This proposal did not simply suggest an operating system change; it introduced a hybrid concept of merging the best of thin-client technology (the simplicity of dumb terminals) with fat-client technology (reliance on a local hard drive for the operating system and application storage) at an affordable cost. This could be accomplished by creating workstations that received a clean hard drive image, free of

corruption, every time the machine starts. Further, we could adapt the system into using its RAM memory as the hard drive, eliminating the need to have physical hard drives on the workstations thereby reducing workstation cost, increasing disk access speed, and eliminating a major source of heat and system failure.

Based on extensive research conducted in our pre-planning phase, the hybrid concept we developed is unique and had not been documented and/or operationalized by any other institution. Because the Linux operating system is open source, (code is readily available, customizable with little or no restriction, and free of charge), we could respond to requests from Library personnel for customizations that would enhance our services. Using a Linux platform would enable the Library to tap into the open source collective of experts, programs, and other resources which would complement the Library's existing expertise, address user needs with fewer barriers towards solutions, and allow for creative endeavors that were not available while using Microsoft Windows. Open source would allow us to modify an application or operating system source code (program) to behave and display exactly what we wanted. In addition to the Linux operating system, there are 90,000 open source applications for our use at any time.

The risks for the Library were high because evidence of application of Linux was scarce, and there were no models from which we could draw. If successful, however, we knew we would eliminate our high-cost reliance on Microsoft architecture, hardware, and software that was consuming ever-increasing portions of our shrinking budget. We would create a platform on which the Library could implement innovative services, sustain infrastructure, and take full advantage of emerging technology.

When our research began in 1999, few institutions had attempted a system-wide migration to Linux. We envisioned paving the way for other organizations to apply the benefits of open source technology as well as contribute the results of our innovation to the library community. The project proposal was named "E3" to signify that this would be the third evolution of the Library's computing architecture.

PLANNING & IMPLEMENTATION

Project Goals and Objectives

The goal of E3 was to develop and implement a new computing architectural model that would be sustainable over the long-term to meet the Library's need for flexible application offerings through code manipulation, increased reliability through the reduction of failure points, and cost containment while eliminating 3-year cycle purchases and MS-related expenses. The objectives listed below would be achieved by migrating both client workstations and servers from Microsoft Windows to Linux:

- Develop a user workstation experience that contributes to individual education and research goals
- Reduce Total Cost of Ownership (TCO) by 50%
- Develop a foundation for long-term technology infrastructure planning
- Provide a long-term technology funding strategy.

Our measures for success include the following:

- Public workstations would provide core Library resource services
- The Total Cost of Ownership would be reduced
- A diskless workstation environment would improve reliability and performance.

Timeline and Steps

In April of 2002, following approval from the Dean and the Library's Steering Committee, TS&D staff began project planning. With Linux identified as a viable alternative operating system, our next step was to verify the availability of applications to replace Internet Explorer and other applications. We identified equivalent applications that could be customized to meet our needs, including Mozilla Web browser, Adobe Acrobat, WINE for Microsoft reader applications, and KDE for a desktop interface. In June 2002, project implementation began.

Hardware requirements were also explored. In order to reduce time devoted to troubleshooting, it was critical to improve the reliability of workstation hardware, network connectivity, and performance. Traditional vendor offerings that met our specifications were cost prohibitive at $1,100 per workstation and $18,000 per server. After conducting a build vs. buy analysis, we determined that we could build workstations at less than $500 per unit and less than $5,000 per server, using readily available, industry standard, off-the-shelf components. TS&D trained student employees using a combination of documentation, DVD motion video, and hands-on training (3 hours total) to enable them to assemble 70 workstations at a rate of 30 per day, resulting in a total of $312 in labor costs. TS&D full-time staff built 14 servers averaging 1 server per day. These machines were assembled "just in time"

throughout the course of the project. Since both workstations and servers did not need to be Microsoft OS dependent, hardware components were selected based on cost, performance, reliability, and availability. Our choices were no longer driven by Microsoft recommendation lists, giving us the freedom to utilize the parts that would best support the project. Building the units resulted in a total savings of $97,683 for the Library.

From November 2002 to June 2003, tech staff used original programming and customization to develop a stable, intuitive, and user-friendly workstation environment that included Web browser, print subsystem, authentication, network file storage, removable media storage, screensaver, disk quota management, multiple/switchable virtual desktops, and disk formatting tools. All were designed and delivered to allow maximum flexibility for the novice user as well as advanced/full features for those with more experience.

In the past, sophisticated users' needs could not be met because meeting those needs would mean compromising the integrity of the workstations. Changes resulting from the E3 project have expanded the scope of offerings available to all users from novice to the "power users" who can exploit the capabilities of the system to its fullest.

From April 2003 to July 2003, we worked to develop a workstation experience that would contribute to the goals of our users. Customer surveys indicate that Fletcher Library users rate accessibility and reliability of public workstations and electronic resources as very important. We approached the user workstation from two fronts: user environment and systems design. The user environment including desktop icons, application "eye candy" modifications, background/welcome screens, etc. were designed by the Library's Public Interface Team which consisted of librarians and public support staff. TS&D staff addressed the technical aspects of making the new user interface a reality. This arrangement allowed us to apply expertise in the appropriate areas and to create an inclusive experience for Library staff.

RESULTS – BY OBJECTIVE

1. Migrate both client workstations and servers from Microsoft Windows to Linux.
In July 2003, all workstations and servers were operational using the Linux operating system and our customized applications, subsystem, and architecture. In August 2003, all workstations and servers were available to our users in a production environment.

2. Develop a user workstation experience that contributes to individual education and research goals.
Our new system provides a level of system reliability which exceeds our previous environment. Users can be confident that their workstation is free of defect and will operate without problems. Our new interface presents library tools to the user based on librarian-guided design and delivery, thus aiding in making the user's library experience productive. Previous fiscal years averaged 372 reported problems. Based on FY 2004 statistics, the new environment generated 32 trouble tickets, a 91% reduction.

3. Reduce Total Cost of Ownership (TCO) by 50%.
The Library used a simplified formula for measuring TCO: TCO=Equipment+Staffing+Incidentals+Training

- Staffing includes all hourly and salary costs associated with staffing the support of the item.
- Equipment includes the cost of purchasing all necessary software, licenses, and hardware.
- Incidentals include peripherals, replacements, and supply costs for maintaining the item.
- Training includes formal training, reading materials, conferences, etc. in support of the item.

OLD DESIGN		NEW DESIGN	
Full-time staff	3 FTE	Full-time staff	2 FTE
Student Staff	1.25 FTE	Student Staff	1 FTE
Workstations	145,600	Workstations	51,140
Servers	36,000	Servers	31,717
Software	8,790	Software	850
Peripherals/Supplies	1,500	Peripherals/Supplies	1,500
Support Material/Training	1,200	Support Material/Training	1,200
TCO/unit	2975.75	TCO/unit	1851.64

Total savings per unit - $1,124.11; TCO savings $134,893 for year 1, non-depreciated
TCO savings assumes staff are dedicated solely to the item measured. *This is not the case and, therefore, realized savings is actually higher*. It is also assumed that Incidentals will not be equal between the two TCO design models. Since TS&D has control over the quality of parts and workmanship in the new model, the new model may result in additional lowering of TCO. TS&D was able to eliminate one FTE Windows hardware/software support position.

4. Develop a foundation for long-term technology infrastructure planning.

Using an open source operating system and open source software that utilizes open systems standards has presented the Library with opportunities to access new and exciting technology developments. Not only does the E3 model break the reliance on a 3-year hardware replacement cycle, but also allows flexibility in systems design to accommodate future needs. There are two key factors with the new architecture: 1) Technology staff easily and inexpensively support and integrate unique makes and models of computers concurrently. 2) Planning can now focus on development of new services and refinement on existing services instead of planning for the next version of Windows, support applications, and managing the resulting problems. By implementing a LAMP-based (Linux, Apache, MySQL, PHP/Perl/Python) architecture, the foundation has been established for Library staff to create new services and methodologies uninhibited by the limitations of a proprietary operating system or restrictive applications.

5. Provide a long-term technology funding strategy.

E3 allows the Library to employ a flexible technology funding strategy. We no longer invest time developing complex 3-year plans, and we can now make technology investments based on available resources. The E3 model also accommodates flexibility in scaling and replacing computing hardware. We can replace a small set of equipment during one fiscal year, and another set of equipment the next fiscal year. Our workstations are no longer dependent on having identical specifications, a major factor with public Microsoft Windows machines. By building our own machines, we have eliminated the need to purchase and maintain vendor warranty contracts, a costly, complex, and time-consuming activity required by our old architecture and system. We can now use "effective lifecycle" as a supporting factor regarding hardware replacement. For example, we can replace system fans every year, power supplies every two years, and CPUs every five years. Historically, monitors have lasted at least six years and thus can be put on an extended lifecycle. Since most of our services are network-centric, we only need to add another network service/server and do not have to replace our workstations to accommodate the service. These strategies relieve the Library from planning for a large expenditure every three years, reducing forced purchasing. We now focus on cost-effective, incremental changes.

MEASUREMENTS

Survey Data

All public workstations allow the user to provide feedback through an electronic survey instrument. For fall 2003, the Library had 78,699 ASU and 9,770 non-ASU accesses accounting for a survey response of 188 users. Sixty-five percent of respondents found the new system comfortable or very comfortable. Only 13% were not comfortable with the new system.

TCO [Total Cost of Ownership]

TCO pre- and post-E3 indicates a 50% reduction, a $134,893 savings.

Strategic Plan 2004-2008

All aspects of the E3 architecture and project are congruent with the Library's Strategic Plan. E3 provides the foundational access support for Library strategic Goal B - to position and promote the library as a place, physical and virtual, that supports teaching, learning, creativity and community; and strategic Goal C - to support scholarship through collections, information access, and content management.

Budget

The E3 model breaks the dependency on a 3-year replacement cycle. The Library can accommodate virtually unlimited hardware types and configurations utilizing the inherent features of Linux and the boot image services. Since TS&D supports hardware replacements in-house, the Library no longer pays for warranty/service agreements. In addition, hardware can be replaced as needed in small or large quantities over short or extended periods of time using available funding and based on need requirements. E3 has created a 50% annual reduction in capital equipment spending to meet future needs.

Trouble Tickets

We have realized a drastic reduction in reported systems and software problems. Previous fiscal years averaged 372 reported problems. Based on FY 2004 statistics, the new environment generated 32 trouble tickets, a 91% reduction. This reduction represents one FTE staff position that was eliminated. We also reduced TS&D student employment by .25 FTE.

INNOVATION DISSEMINATION

To date, over 1200 institutions, businesses and individuals have downloaded the project documentation and applications from our Website. We have also received acknowledgement from many libraries, academic institutions, and library-related companies on the success of this migration including: Notre Dame Libraries; University of Angers (France); Whittier, CA Public Library; Oxford University; Michigan State University; National University of Cordoba Argentina; University of Connecticut Libraries; BirdLife Indonesia Programme; Fu-jen Catholic University Library, Taiwan; University of Arizona Library; Midlands, TX Technical College; Amigos Library Services; and Arizona State University Information Technology. This list is not inclusive but does indicate the breadth of interest from the local, national, and international communities. Our project has been cited by Purdue's Open Source Development and Documentation Project, a presentation to the Council on Computing at Dartmouth, an article on open source solutions by the School of Library and Information Sciences at San Jose State University, a bibliography at Victoria University of Wellington, New Zealand, and in a presentation at the 2003 French Libraries conference.

Library and open source related listgroups were notified at the start of the project and provided updates. Article submission queries were initiated to library and academic journals including Computers in Libraries, Electronic Librarian, Library Information Technology Association Journal, Syllabus, and EduCAUSE. Announcements were made on websites including linuxpr.com, oss4lib.org, and lisnews.com. Presentations detailing the E3 project were made at Syllabus 2004 in San Francisco, CA, Library Information Technology Association National Forum 2004 in St. Louis, MO, and the upcoming Computers in Libraries 2005 conference in Washington, D.C.

Throughout the project, activity reports and detailed documentation totaling 184 pages were provided through various communication outlets. We provided West campus IT with information about the project and sent progress reports. In December 2002, TS&D initiated the Linux User Group on the West campus to facilitate sharing of knowledge among the ASU community. The ASU University Library Instruction Systems and Technology (LIST) group was given the project plan as well as updates to inform their staff. Bill Lewis, ASU Chief Information Officer and Vice Provost for IT, was also informed about the project and its outcomes.

The Library has an informational Website (wlibinfo.westlib.asu.edu) that makes available for public download all custom applications and processes created by TS&D for the E3 model. This site is also listed at Google. The Fletcher Library Website (library.west.asu.edu/aboutus/index.html) provides a link to the complete project description and all current documentation for the general public to view.

CONCLUSION

The implementation of E3 has propelled Fletcher Library into the forefront of library technology applications. The most important result is our increased capacity to be responsive to the technology and research needs of our users with the resulting impact on their studies and research. Equally important, we will be more effective stewards of our equipment and time, implementing new technology that will not adversely affect our budget.

TS&D has created a technology roadmap propelling E3 forward through 2007. We have developed and are using version 2.0 of a public machine image and will implement an E3 portal server for information sharing and development among other libraries, with the project and information exchange starting in the summer of 2005. Development of administration tools will begin in fall 2005 to simplify administration and application of the E3 system by other organizations. In 2006, a major change in the overlying desktop environment will start with the implementation of the Sun Looking Glass 3-dimensional desktop environment with a deployment target date of 2007. These plans are intended to enhance the user experience, further cost-containment, and maintain current applied computing technologies.