

Managing the “Forest” of Administrative Servers and Documents

By Terence K. Huwe

The University of California, Berkeley provides a good case study for thinking about—well, a whole lot of things, but certainly it's on the front lines of the crisis in computer security. Berkeley is a frequent target of hackers, and they have a lot of targets to choose from. As a key contributor to the Internet backbone, as the cradle of the Unix C Shell, as the birthplace of the Free Speech Movement—well, who among hackers wouldn't notice the plum potential. What's more, server administration has been decentralized for years, creating an environment with thousands of servers running on multiple platforms, including several flavors of Unix, every kind of Microsoft product, and a very healthy Apple presence. The security statistics are alarming: security “events” average 25 per day, 3,000 messages including viruses arrive daily, and “spams” number in the tens of thousands, according to Jack McCredie, the campus Chief Information Officer.

Office culture plays a role, too. Decentralized systems administration reflects the campus bias for decentralized teaching and research. This is important, because it takes a certain amount of organization-wide planning to sustain security, and that means involving lots of staff, not just technologists. Decentralized server administration provides autonomy, but it also brings security risks right into the lap of local administrators. Which means hijacked servers that are used to nefarious ends.

It's not hard to find hacking horror stories. Last year the Institute of Industrial Relations' main file server was hijacked after a sustained attack that lasted weeks, and had to be shut down and cleaned. The University Library's servers were seriously assaulted some years ago, and restoration required a crisis response that lasted three days. And then there's those puckish “inside” jobs. Our first file server was configured by my student webmaster in 1996. He created “back door” modem access to the server, to avoid the queue for busy student modems. He also shared it with friends (until we found him out.) Later, when he went to work for Hotmail, just before it was bought by Microsoft, he finally noticed that we had shut it down. We had a laugh about it. Of course, by then he made his first five million, at the age of 20. Oops, sorry—that was another era, I guess.

Organization-wide responses to security threats—or the lack of them—have a big impact on digital libraries. Our “parent” organizations—colleges, cities, and firms—often define technological platforms without much thought about library requirements. This makes it all the more vital to make sure our voice is heard. Even then, we don’t always get to choose how things are configured. But keeping track of the planning process helps us take advantage of new initiatives at the earliest possible moment.

Those of you who work in single platform or unified computing environments may shudder to hear how vastly diverse campus computing is at Berkeley, but it has carried advantages. It’s hard to herd cats and faculty members, so the campus tends to adopt new platforms at the end of the product cycle rather than the beginning, and collective knowledge has grown in the mean time. One of those moments popped up at Berkeley just a few months ago, and its pilot phase is far enough along for me to offer a note of optimism in the hacker wars. Oddly enough for Unix-centric Berkeley, a Microsoft application provided a nudge that created critical mass.

The Active Directory “Forest”

Microsoft Windows 2000 Active Directory (AD) has many enterprise-level applications, including security. AD recently gained the attention of central campus computing professionals who were juggling a major e-commerce portal rollout, a “single logon” smartcard known as “CalNetID”, and a vast problem in sustaining decentralized servers that had sprung up everywhere. Each of these large-scale projects involved big changes in the campus computing culture, and security issues lay at the heart of the various initiatives. With over 2,000 Windows servers and 60,000 CalNetID accounts in use, campus planners determined that AD architecture was a good management tool. AD architecture offers a centralized, secure domain that not only improves security, but also eases overall systems administration. The local name for the pilot version of the Active Directory environment is “the Forest” (<http://istpub.berkeley.edu:4201/bcc/Winter2002/infr.calnetad.html>.) The Haas School of Business and the Institute of Industrial Relations were both early adopters of the pilot plan.

Microsoft describes AD as a means to store information about network objects. It makes object information available and usable to users, computers, and applications dynamically. AD uses common standard technologies like Lightweight Directory Access Protocol 3 (LDAP3), Domain Name System

(DNS), and Kerberos 5 authentication. In plain English, it enables administrators to take dynamic hosting to a new level, managing whole directory trees of files and applications in a single container. This capability enables the central campus computing authorities to offer departments a much more secure environment and a chance to retire some servers, while preserving all of the local naming and domain conventions that have grown like topsy around the campus.

That's a big deal politically, because Web names and presences reflect real-life egos, not to mention the loving handiwork of hundreds of systems administrators who have big-time control issues. "The Forest" is exactly what it sounds like: groves of directory trees, living independent lives under the watchful eyes of forest managers, who had a new incentive to partner with department-level server managers. Existing folders and directory trees continue, but need not exist on local servers any longer. Everyone gets to be who they always were on the Internet, but they gain redundancy and higher security. The catch is that they will also find themselves conforming to a campus-wide standard that will slowly, but surely reduce the extreme diversity in server configuration.

Join, or Die

The Forest is an interesting example of what's possible when everyone agrees to work together, emphasizing "carrots" instead of "sticks". Moreover, it's a good example of how to approach a problem about which we think we know a lot—security—from outside the box. In the firewall-free zone of academic computing, security breaches fall into three really big categories: outside hackers, inside pranksters, and freeloaders who want free access. In the absence of firewalls, IP authentication creates the de facto "extranet" of campus computing. But it's only as safe as a porous culture can make it. The Forest offers Berkeley departments better safety from each of these hacker blocs, even though its primary goal is administrative. Nor does it call upon anyone to change their directories or political stance on open source computing—they get to keep everything. That spells opportunity for digital librarians with persistence on their minds.

Three Crossover Applications

As a profession, we're better than most at identifying long-lasting value in really obscure sets of data and files. However, I think we could do a better job of monitoring how administrative computing can favorably impact digital librarianship. The line between "collections" and administrative computing is

getting pretty fuzzy, except where standards are spelled out in the extreme, as is the case with online archives that utilize the Dublin Core Metadata. AD projects like the Forest suggest at least three potential opportunities for digital libraries that flow from the security-driven initiative of the Forest.

First, fewer systems administrators means more money for content. Personnel costs are always the biggest piece of the pie. Berkeley is a community of 30,000 students and 12,000 faculty and staff, and there are lots of programmers. AD initiatives like the Forest will create moments in time when this workforce can be shrunk back to its pre-Internet size, through attrition, without a reduction in service. As budgets shift, it will be up to information specialists to capture their fair share of the released funds.

Second, some uniformity is a good thing. A phenomenal amount of intelligence has been thrown at the Web sites that we now depend on for lots of tasks. Academic Web sites are often among the best, because there's less market hype and more long-term stability. Libraries, particularly small ones, can reap a big benefit first by cleaving to enterprise-level standards—and then by shaping them to benefit the needs of library collections and metadata.

Third, less time spent on server administration and security threats means more time to build digital collections. In these times, even smaller libraries often employ a systems administrator, or recruit someone who can juggle systems responsibility with related skills, like electronic outreach and training. Probably most of us are in that second boat, and any time saved from the mundane tasks of server administration is welcome news. At the same time, the fuzzy line between “content” and administrative files invites us to increase our roles as records managers. Systems librarians may find themselves in the position of being frontline collaborators with archivists—if they aren't already playing that role.

Breaking the Cycle of Conventional Thinking

While I've been weaving together the threads of server administration and computer security, it occurs to me that the most interesting aspect of the Forest is its collaborative aspects. Even though I favor administrative decentralization from an intellectual viewpoint, decentralization fosters more moments when security can slip. In the local culture at Berkeley, it's possible for Berkeley programmers to meet very few of their colleagues over the course of a career. They often make significant purchases or strategic decisions without consulting others. The Forest, with its emphasis on Microsoft Active Directory architecture,

creates a real incentive to retire department-level servers (and eliminate their long term workflows.) It enables financially pressed departments to de-emphasize the “turf value” of local control of technology – and spend their funds on their academic mission.

Central campus planners broke away from conventional thinking by using this newish technology as a “carrot”, when they might have made it into a “stick” by charging for it. Librarians, who compulsively want to give everyone information free of charge, are far ahead in understanding how to broker collaborative relationships around technology. Whether in the arenas of security, authentication, or records management, we have an example to make – if we are bold enough to take that step.