

The Systems Librarian

by Marshall Breeding

Offering Remote Access to Restricted Resources

Authentication, authorization, and proxy are the keys to success

At Vanderbilt—as I’m sure it is at other libraries—we focus much of our attention on finding ways to provide remote services to patrons who don’t physically visit the library. Last month, I described some of the issues and options related to providing reference services to remote users. In this month’s column, I’ll discuss some of the issues associated with connecting remote users to those electronic resources that must be restricted to the library’s own clientele.

Libraries have a growing base of users who want to take advantage of our information resources from their homes, offices, and dorm rooms, or from any part of the world as they travel. One of the key challenges to providing services to these patrons involves identifying them as valid library clientele and enabling them to gain access to licensed resources that aren’t part of the “free” Web. In more technical jargon, in order to offer access to these services, we must deal with issues of authentication, authorization, and proxy.

Libraries routinely license content from publishers and database producers for significant fees. These resources are often not freely available to anyone on the Web, but only to those who have valid subscriptions. Publishers implement various mechanisms that block access to those who aren’t known subscribers. The most common of these methods involves IP address restrictions.

Restriction by IP Address

All computers connected to the Web have an IP address. We’re all used to seeing these addresses written as four decimal numbers separated by periods. The IP address of my computer, for example, is 129.59.150.100. Publishers have the ability to control who may use their products by maintaining access lists of allowed IP addresses. When someone attempts to get into an IP-restricted resource, the system simply looks up his or her IP address. If it finds the address in the list, the individual is allowed in. Otherwise, the user is rejected.

It’s not always necessary to deal with IP addresses one by one. These access lists also use network address ranges. IP addresses are organized in a way that easily identifies whole networks of computers. All computers on our main campus, for example, begin with 129.59. This way, we can use the abbreviation 129.59.*.* to represent all the computers on the main part of our network. When purchasing content from an information provider, the usual procedure involves submitting a list of IP addresses or IP networks so that it can be added to the access-control list for that subscribed resource.

I’ve been following authentication and authorization for a long time. For at least the last 5 years, there has been much discussion that IP authentication isn’t sustainable and will soon be superseded. Many have said (myself included) that more sophisticated technologies—such as Public Key Infrastructure (PKI)—will surely emerge to re-

place primitive and cumbersome IP authentication. However, there are few signs of any movement away from IP authentication. While more advanced authentication systems like PKI offer a more robust authentication environment, implementing and managing these new systems is complex and expensive.

So, it appears that IP restrictions may continue to be the common approach used by libraries and information providers for quite some time. We’d do well to find methods to make this approach work in the best way possible.

“
*Ideally,
a proxy service
would automatically
step in when
it’s needed and
wouldn’t require
users to change
their Web
browser settings.*
”

Problems with IP Authentication

On one level, this arrangement works well enough. When we purchase a subscription to a resource, we give the provider a list of valid IP addresses to place in its access list. Users who attempt to access the service get through as long as their computer falls within the set of allowed IP addresses.

The fundamental problem with this scheme is that there are many circumstances in which library users work from computers that aren’t associated with the organization that sponsors their access to licensed content. When people use the Internet from their homes or offices, they’re likely getting that access through an arrangement with an Internet Service Provider (ISP). Their computer’s IP address falls within a range associated with the ISP. When connecting to the Internet through an ISP, the computer has no guarantee of getting the same IP address for each session. Since ISPs serve a wide range of customers, there’s generally no way to correlate the IP addresses of a given individual with his or her library.

In the academic environment, most universities and colleges are moving away from providing their own dial-in connections to their networks. With dial-in services, the users who connect from their homes and offices receive IP addresses that

are part of the university’s IP addressing schema. But operating a dial-in service is expensive and requires constant upgrades—especially as modem technologies evolve. ISPs are generally in a better position than universities to offer these services. The advent of high-bandwidth Internet access through cable modem service and DSL technologies makes ISP access to the Internet even more appealing than institution-provided service.

Proxy Servers Fill the Gap

So, given that fewer and fewer of our remote users have IP addresses that we can identify—and given the reality that information producers continue to rely on IP addresses for access—what can a library do to facilitate access to these services for our remote users? The approach that seems to work best involves the use of a proxy server.

Proxy servers, at least in this context, allow a remote computer to access IP-restricted resources. This works by altering the remote computer’s network traffic in such a way that it appears to any IP-restricted service as having an IP address that falls within the allowed range.

Authentication

Before a proxy server will perform this task, it has to be sure that the user is associated with the institution. Most organizations have some sort of username and password scheme to identify their patrons. A university might have a campuswide authentication system used for e-mail and other network services. These schemes may use Kerberos, Radius, or Windows NT domain servers for the actual authentication routines. A library automation system can, in some cases, be used as an authentication server. Many libraries have automation systems that maintain usernames and passwords for renewing books, paying fines, or making ILL requests. Innovative Interfaces, Inc. offers a module for its INNOPAC service called Web Access Protocol that can be used for this purpose. In order to make a proxy service work for controlling access to library applications, some type of username/password authentication scheme is needed.

The traditional proxy server model requires users to adjust the setting in their Web browsers. They must program their browsers in advance with the proxy server’s IP address and port settings. Once set and activated, the browser passes its traffic through the proxy server. To initiate the session, users are prompted for their username and password. The proxy should be activated only for access to restricted resources. With most Web browsers, enabling and disabling proxy services is a manual process that takes a few steps.

Problems with Proxy Servers

This model of using a standard Web proxy server was how we initially provided authenticated remote access to our restricted electronic resources. For the most

part, this method worked, but there were some significant problems. While experienced Internet users had no trouble following the few steps required to set up their Web browsers to use the proxy service, there were some who found the process too complex. The steps made the settings vary between Netscape Navigator and Microsoft Internet Explorer and among version numbers—and even vary according to whether the service was running on a Macintosh or a PC. Writing simple, easy-to-understand instructions that covered each of these permutations was a challenge. Another problem we encountered was that some users would set the proxy settings and leave them enabled permanently. This doesn’t necessarily cause problems for the users, but does mean that all their Web surfing is channeled through our proxy server, adding to the overall load.

But the most difficult problem arises when a user needs to interact with multiple proxy servers simultaneously. We found that some of the major ISPs in our area require that their customers go through a proxy server maintained by the ISP. Disabling this proxy setting would cause problems with services offered by the ISP. It was especially complex for the user to have to use one set of proxy settings for library resources and another set for ISP-related services.

EZproxy to the Rescue

The problems we encountered are typical of the traditional Web proxy server. Fortunately, there are other approaches. Ideally, a proxy service would automatically step in when it’s needed and wouldn’t require users to change their Web browser settings.

“
*The main advantage
of using EZproxy is
that library users
don’t have to make
any changes to their
browsers to access
restricted resources.*
”

After some investigating, we found all of those characteristics in a product called EZproxy from Useful Utilities of Glendale, Arizona. EZproxy is a URL-rewriting proxy server that offers an approach that’s well-suited to the problem of providing remote access to library resources. It operates between the user’s Web browser and the restricted service, intercepting both the Web browser’s requests and the pages returned by the Web server. From the perspective of the restricted Web service, the requests come from an authorized IP address and are accepted as valid. All the URLs on the Web pages that returned from the restricted service are rewritten as they’re displayed on

the user's Web browser with the base address of the EZproxy server. This rewriting is necessary so that the links will function correctly—passing through the proxy server—if they are clicked on.

With EZproxy, the initial links to the restricted resources must be written in such a way that they go through a script that directs them to the EZproxy host. For example, to connect to ProQuest—one of our most popular restricted resources—the patron uses a link (<http://proxy.library.vanderbilt.edu:2048/login?url=http://www.umi.com/pqdauto>) instead of the native URL (<http://www.umi.com/pqdauto>).

In addition to using the new form of the URL for links to the restricted resource, an entry needs to be made in the EZproxy configuration file.

Once the resource has been properly configured, users will be able to access it easily. If they access it from an address that's within

the library's IP range, they are taken directly to the resource. When it's used from outside the library's IP-address range, EZproxy will then prompt them for their username and password. If entered correctly, they are passed on to the resource. EZproxy will maintain the connection, adjusting all the Web pages so that they'll work correctly.

As EZproxy provides access, it logs each use into a file that can be utilized to create statistical reports. These statistics contain information that allows the library to monitor the use of specific electronic resources. This information is especially helpful in making renewal or cancellation decisions for these resources.

The main advantage of using EZproxy is that library users don't have to make any changes to their browsers to access restricted resources. All they need to do is provide their username and password when asked. While the library must perform some up-front work,

it's well worth the effort in order to create a more user-friendly environment.

EZproxy is very inexpensive, with a list price of \$495 for a one-server license. It will operate under either UNIX or Windows NT. The supported authentication methods include Radius, LDAP, UNIX login accounts, the INNOPAC Patron API, or a local file of usernames and passwords.

Automating the Library's Work

As I noted above, to use EZproxy the restricted resources' URLs have to be adjusted. That part of our library's Web site is driven by a database in which the pages are generated dynamically. For each resource the library subscribes to, there's a database record that includes its name, URL, description, subject fields, and other metadata. Since the pages that provide access to these resources are already generated dynamically, it was easy to enhance the script

to alter the URL to the form needed by EZproxy. We created these dynamic Web pages to provide easy access to the library's electronic resources, including the ability to find resources by subject categories, by keyword searches of titles and descriptions, and through browsing alphabetical lists. The ability to seamlessly integrate EZproxy for remote users was an added benefit.

If you're interested in looking at our database-driven access to electronic resources, go to <http://www.library.vanderbilt.edu/heard/edatabases.shtml>.

For more information on EZproxy, see <http://www.usefulutilities.com/ezproxy>.

Marshall Breeding is the technology analyst at Vanderbilt University's Heard Library and a writer and speaker on library technology issues. His e-mail address is breeding@library.vanderbilt.edu.

Endeavor Announces Commercial Release of ENCompass, Deployment of Voyager in Ithaca College Library

Endeavor Information Systems has announced the first commercial shipments of the ENCompass digital management system to three additional customer sites around the country. Endeavor has also announced that the Ithaca College Library in Ithaca, New York, has chosen the company's Voyager integrated library management system.

ENCompass

Already in place at three development partner sites, ENCompass now provides even more academic and research libraries around the world with integrated access to their print, electronic, and digital collections, including local and remote electronic databases and local digital collections. Designed for unified access to the varied collections of any library, ENCompass allows libraries to design master architectures of print, electronic, and digital collections and offers users a single search across all formats with relevance-ranked results sets. The only product of its kind on the market today, ENCompass provides the creation

of metadata, the organization of collections of metadata, and an overarching layer of navigation to assist researchers in discovering previously underused digital collections, according to the company.

ENCompass continues to evolve with the assistance of three development partner sites: Cornell University, the Getty Research Institute, and Kansas State University. Each site is unique, with private and public collections integrated into ENCompass at the institutions. According to the announcement, ENCompass allows each institution to organize digital collections in ways that make sense to the resources and needs of that institution. Cross-institution collection development can also be a reality with the integration of Z39.50 in ENCompass.

For staff, ENCompass provides a simple Windows-based drag-and-drop collection-building feature with reduced data-entry-organizing digital collections. For navigation, ENCompass includes a powerful search engine that grasps all metadata forms, regardless of digital format.

Ithaca College Library

Serving over 6,000 students and 500 faculty members, the Ithaca College Library plans to implement the Voyager system this September. The library's Voyager launch will be simultaneous with a major renovation of three floors of the library.

With so many projects on the table, the library staff has a plan for learning the new system and experiencing a redesigned work environment. "We plan to take our time and use an out-of-the-box version of the Voyager Web catalog for a time," said Margaret Johnson, college librarian at Ithaca College. "The simplicity of Voyager setup and the opportunity to continually modify the Voyager system are benefits for the library staff. Because we can change the system over time to meet our needs, there is less pressure on [us]. Everyone has an opportunity to participate in the ideas for the new look and feel, and this gives the staff's ideas more of a chance to be heard.

"The Voyager implementation gives us an opportunity to re-examine the way we're doing Web pages and ask ourselves

what makes the most sense now, especially with the opportunity for Z39.50 integration of resources," Johnson said. "Now we can evaluate if abstract-and-indexing databases belong in the catalog instead of where we currently display them, on a separate page of our library Web site. There are lots of options and chances for combinations of resources...."

Founded as the Ithaca Conservatory of Music in 1892, Ithaca College continues to uphold a music-study curriculum in addition to a broad range of studies in humanities and sciences. The college's music focus is supplemented by a large music collection in the library, including scores, sound recordings, and print materials. In addition to the Milton Cross Collection of radio scripts and memorabilia, special collections at the Ithaca College Library include the Rod Serling Archives, a grouping of television scripts, films, and related materials.

Source: Endeavor Information Systems, Des Plaines, IL, 847/296-2200; <http://www.endinfosys.com>.

Inmagic

(continued from page 51)

Version 2.2 provides librarians with new levels of control over numerous customization features. "We've enabled a comprehensive suite of customization tools throughout the integrated Windows and Web interfaces," said Moulton. "Librarians need control over the look and feel of their end-user's experiences. Version 2.2 greatly enhances their ability to do this by allowing them to customize the fonts, colors, defaults, etc., exactly the way they want them to be."

Information-request forms let users submit reference queries via the Web to library staff, where they can be tracked and reported on throughout the life of the query. When a user makes a reference request, an e-mail is sent to library staff. Simultaneously, the query also goes into the catalog

as a work-flow item. As progress is made on meeting the user's request, BiblioTech PRO's Record Update module can be utilized to update query status. At any time, BiblioTech PRO's Report module can report on work-flow history and current status.

Additional enhancements in version 2.2 include automatic spell-checking, the ability to define multiple URLs or image (documents) files for each citation, and the ability to save Web search results to a Word or text file.

DB/TextWorks v5.0

DB/TextWorks v5.0, which includes the Web publishing components DB/Text WebPublisher and WebPublisher Lite, now features relevance ranking as well as other enhancements.

"Relevance ranking allows users to spend less time searching for information,"

said Phillip Green, Inmagic's CEO. "It gives them the ability to sort reports so that records most similar to search criteria appear first. This creates huge savings in time and efficiency for our customers."

Graham Beastall of Soutron, Ltd., an Inmagic dealer based in the U.K., said, "Version 5.0 of DB/TextWorks gives librarians the ability to provide their users with simple and refined searching similar to that provided to users by systems costing many times that of DB/TextWorks."

Relevance ranking is not the only new feature enabled through version 5.0. While performing database maintenance activities (such as checking a database for errors), users can now simultaneously rebuild indexes when using version 5.0. This enhanced functionality allows users to perform database maintenance activities more easily, with less time expended.

The product's e-mail capabilities have also been expanded, creating easier, less time-consuming, and more powerful experiences for users. Reports sent via e-mail can now include attachments. In addition, user-defined scripts can now send e-mail without requiring parameters to be supplied by the user. This saves time and increases accuracy during repetitive e-mail tasks.

Enhanced display features let users in DB/Text WebPublisher or WebPublisher Lite review detailed and summary information at the same time. Expanded record displays can now appear in separate browser windows, giving users the option to point, click, and manipulate information with more flexibility.

DB/TextWorks v5.0 is available to those who have enrolled in Inmagic's new Inmagic-ADVANTAGE maintenance program.

Source: Inmagic, Inc., Woburn, MA, 800/229-8398; <http://www.inmagic.com>.