Shibboleth & Access to Licensed Content

Mark Earnest
Lead Systems Programmer
The Pennsylvania State University
Current/Future Resources

- Napster
- JSTOR
- OCLC
- Elsevier
- ProQuest
Issues

- Access to library resources currently controlled on a hybrid IP Address/Access Account basis
- Some licenses dependent on campus location
- To wayf or not to wayf...
- Deep Linking
- “Special” Cases
Napster

- Shibboleth used for student registration
- Population limited to Residence Hall students
- No identifying attributes to be sent
- eduPersonTargetedID used for persistence
Library Resources Today

- Starts with the CGI
  - If the IP address is allowed to access resource, the request is passed along
  - If not, the user is redirected to EZProxy
  - EZProxy calls a custom exit we have written to authenticate the user via their AccessID/Password (Kerberos) before allowing them through to the Resource
EZProxy

- Popular software to act as a intermediary between IP restricted content and remote users
- Rewrites web pages as they are returned to the user so that site links go through EZProxy
- Provides “exits” to call authentication routines, but does not work with Apache authentication modules.
EZProxy Configuration

From ezproxy.usr file:

- ezuser:ezpass:cgi=https://sserver4.lcs.psu.edu/scripts/authdisla.exe?

- ezuser/ezpass - EZProxy userID and password

- The CGI is called with the url of the requested resource as a parameter.

- The CGI authenticates the user, then uses the EZProxy uid/pw to obtain a session ID, which it passes on a redirect to EZproxy (along with the URL or the resource)

- EZProxy sets the session ID as a cookie in the user’s browser, then obtains the required library resource they requested.
PSU’s EZProxy Plans

- Rewrite our CGI to point directly to Vendor’s Shibboleth Target sites (if applicable), bypassing EZProxy

- Write an EZProxy authentication module to work with our University’s web single sign on solution (Cosign), which is also what our Origin Site will use.

- This will give us web single sign on regardless of how the content is accessed.
Shibbifying EZProxy

- For universities using Shibboleth as web single sign on, or just desiring a single sign on to library resources.

- Scott has done this by pointing EZProxy’s external auth hook to a perl script protected by Shib (similar to what we plan to do with our Web SSO)

- Is there a better way?
JSTOR

- We have successfully used Shibboleth with JSTOR in the past (Version 0.8) and are revisiting them with version 1.2.

- Simple. We license JSTOR resources for all students, regardless of campus or status. No identifying attributes need to be sent.

- JSTOR is working on the “deep linking” problem, we are watching this closely.
OCLC

- Licensed to all non Hershey students.

- Dual “AuthO” Values presented a unique problem to PSU. Tentative solution is to pass both and let OCLC sort it out.

- AuthO values will be automatically placed in any non-Hershey student’s LDAP entry.
With Shibboleth 1.2 the option is available to use a custom ResourceMapProvider to map the AuthOs to a specific application ID.

The Target can now define an application ID which the ARP can use to determine which AuthO to send:

```xml
<Application id="pluto" providerId="urn:mace:inqueue:brown.edu"/>
```
Access Restrictions

- OCLC is the only Shib enabled resources that has location specific restrictions in the agreement with PSU (So far...)

- Two options, pass location to target or create entitlement value locally based on location.

- Surprisingly (to me anyway) target sites seem content to let us apply the contract rules and make the authorization decision.
Access Restrictions

- As more resources embrace Shibboleth, the access control policies will become more complicated.
  - Some of our vendors go by College, status within the University, Department, etc.
  - Which attributes take precedence?
  - Policy problem, not technical
TargetedID

- Persistent (but opaque) attribute to anonymously allow target sites to “remember” users.

- The value is a hash of the UserID, the Target ID, and a secret “seed” value generated at the Origin.

- Can be generated dynamically or mass generated and stored in LDAP

- Elsevier now supports this for personalization & Napster requires it.
“Gotchas”

- Some LDAP data is not as accurate as we thought

- Collected from multiple sources, never really tested until Shibboleth came along

- There are always strange exceptions.

- Hershey students that also teach at another campus, or vice versa. Many professors and students hold “dual citizenship” at PSU Campuses
“Gotchas”

- Some contracts do not lend themselves to attribute based authorization.

- OCLC AuthO values were only different based on which OCLC resource was accessed, not anything specific to the user.

- Our policy is to provide access to Library Resources to anyone physically in the library

- Continue IP filtering? Use temporary accounts?
Questions/Comments?

- Obligatory final slide
- Thank you :)
- email: mxe20@psu.edu, llg5@psu.edu