

SDSS briefing note — EZproxy Central

Interim report: 14 March 2008

1 Scope

In managing the transition to the UK federation, it had been envisaged that access to resources which support Athens but not Shibboleth could be enabled by the Shibboleth to Athens gateway. This would be made available to the community free of charge until 2011, under an agreement between JISC and Eduserv. In the event, it was not possible to reach agreement, and institutions are therefore faced with unexpected subscription costs if they wish to continue to use the gateway after July 2008.

JISC has advised institutions that would like to participate in the UK federation by choosing one of the two open source implementation options (e.g. full in-house implementation or in-house implementation with paid-for support) to seek an alternative access route for resources that will not be federation-compliant by the end of July 2008.

Many resources also permit access based on IP address checking, which provides a basic solution for on-campus use (though without user personalisation). It is also possible to provide off-campus access to IP-authenticated resources by using a proxy server. Therefore, institutions may consider using a proxy server (e.g. EZproxy) to provide temporary off-site access to resources that are not yet available via the UK federation.

JISC has been asked by several institutions to explore the possibility of providing a centralised proxy service.

2 General approach

The approach considered here is based on EZproxy, a commercial product now owned by OCLC, which is used by university libraries as a means of providing off-campus users with access to licensed resources. Many resources support IP address checking as an access control method; this permits access for users present on the institution's 'secure network'. The off-campus user can log in to the proxy server (using Shibboleth, in our case) and access licensed resources via the proxy. All traffic between the user's browser and the resource travels through the proxy, which makes the user appear present on the secure network and therefore satisfies the service provider's IP address check. The proxy is presented to users as a library portal with links to the available resources; the user must use these links to ensure that the web traffic is routed through the proxy, and so appears as on-campus traffic.

While this is a solution currently available to institutions, it is unlikely to become widely deployed within a short time; institutions would have to acquire knowledge on how to deploy and operate the EZproxy server, design and implement a way of presenting this to its users (the Librarian's responsibility), and provide documentation, user education and support.

The EZproxy Central proposal attempts to short-circuit the need for distributed deployment by providing a central facility available to all institutions. This would provide institutions with a basic form of proxy service, without institutional customisation, but also without the overhead of local deployment.

Proxy technology is well-understood, and the EZproxy product is reasonably regarded. As an engineering exercise, the deployment of a solution based on a central EZproxy service should be technically feasible.

3 Deployment options

The options described below in outline are all variations on a theme. They offer different technical means of providing a central EZproxy facility which acts on behalf of a number of institutions. The next step would be

to document the development activities required in more detail, broken down to institution, service provider and proxy provider levels. No firm conclusions should be drawn at this stage.

3.1 Option 1 - Authority proxy (A-P)

Scenario: Big central box with per-institution licensing tables.

The user accesses the A-P and selects a service from a list maintained by it (using information extracted from the licensing tables). The user is then asked to go through a standard Shibboleth logon with his/her home institution. On successful authentication, the A-P uses EZproxy's authorisation mechanism to determine whether the user's institution is licensed for that service. If so, the service request is routed to the target service.

The service believes this is a licensed user because the A-P is always believed; there are no means of telling the service which institution the user belongs to.

Pros

This has the advantage that it makes use of EZproxy's built-in authorisation mechanism; scripts can be written and applied statically to the raw licensing data to derive configuration language statements that control the authorisation decision. Little technical development is required.

The set-up cost for institutions is negligible, and the configuration and maintenance costs for service providers modest.

Cons

Licensing tables provide a matrix of services against institutions, indicating the services for which each institution holds a licence. This requires every service provider to disclose to the proxy operator its list of licensed institutions (per resource), and agree to inform the operator of any additions or deletions from the list. If the information is considered to be commercially sensitive, the service provider must be confident that the proxy operator will use it only for the purpose stated and that it will not be shared with any third party. This might preclude the use of the licensing tables for other potentially useful purposes such as OpenURL resolution.

The service provider is also required to accept the proxy's assertions concerning the licence status of individual users and to provide access purely on the proxy's say so. This is a substantial leap of faith.

Neither the service provider nor the client institution can be provided with detailed usage information, since there are no means of conveying the identity of the user's institution in each service request. The proxy itself could maintain counts of access requests, but this information is not definitive as a record of end-user activity, which only the service provider can acquire. The service provider could attempt to correlate log data from the proxy indicating the institutional affiliation of each user with its own usage statistics, but this may be difficult in practice.

3.2 Option 2 - Multi-IP proxy (M-P)

Scenario: Big central box; no licensing tables; registered IP address for each client institution.

The user accesses the M-P and selects a service from a list maintained by it; the user is then asked to go through a standard Shibboleth logon. The M-P looks up the user's institution in its database and extracts the corresponding registered IP address. The M-P passes the service request to an EZproxy instance along with the IP address which it is required to present as its source address when acting on behalf of this user.

The service believes that this user is a member of a given institution because the source address has been registered with it as a known IP address for the institution. The service inspects its local licensing tables to determine whether the institution is licensed for the resource and grants or denies access accordingly.

Pros:

The setup costs for both the institution and the service provider are modest. The institution requests an IP address from the EZproxy Central service which it registers with each of its service providers as an additional IP address to be regarded as part of its 'secure network'. The service provider configures the address in its tables as per routine.

Cons:

This option would require substantial development effort since the multiple instances approach is not an inherent capability of EZproxy software. It is possible that OCLC as owners of EZproxy could be persuaded to undertake this work, but the timescale and costs for doing so are unknown. As this development is unlikely to be central to OCLC's business, implementation is unlikely to be seen as a priority.

The service provider would be obliged to register the name and URL of each of its services to enable the M-P to construct a list of available services for presentation to the user. This would require ongoing maintenance. The service provider must also be prepared to accept that the EZproxy facility is performing correctly, and has ensured that only appropriate use is being made of the IP address registered for each institution.

3.3 Option 3 - Virtualised proxy (V-P)

Scenario: As above, a big central box; no licensing tables; registered IP address for each client institution.

This option aims to reduce the development cost of the M-P solution by using the big central box to support a number of virtual machines, each of which acts on behalf of one institution, and uses the IP address registered to the institution in each case. This operates as option 2, except that the per-institution EZproxy instances are realised as virtual machines rather than depending on virtualisation developed in software.

The pros and cons are the same as for option 2, except that the development time should be reduced and would not rely on enhancement of EZproxy by OCLC. Development of a virtual machine capability on this scale is unfamiliar territory, however, and would inevitably encounter unexpected problems. One option would be to contract the task to a company with expertise of this type of development.

4 Cost/benefit balance

The key benefit can be measured by the number of institutions that are satisfied by the Shibboleth/ EZproxy Central solution and do not feel it necessary to subscribe to OpenAthens in order to access non-Shibbolised resources. The proxy service provides only a partial solution (notably, lack of personalisation and user accountability). Consequently, an institution may opt to use OpenAthens if even a single service it regards as critical is not available to it, in fully functional form, via the proxy.

The main costs are likely to be equipment and licensing. Both are tied in with service load, which cannot be predicted with any real confidence. There is no attempt here to distinguish setup and ongoing costs.

- a) *Equipment.* All of the options will require a substantial hardware resource in the form of the 'big thrumming box', or farm of servers.
- b) *Staffing.* This should include coverage for software maintenance and development; administration; statistics processing; liaison with institutions; service negotiation; support.
- c) *Facilities management.* The communications load is potentially substantial, given that all traffic between the browser and the web resource would go through the proxy; it would be unwise to attempt to route this nationally distributed network traffic through an internal institutional network. Colocated servers should be sited externally on SuperJANET.
- d) *Licensing.* A significant number of institutional EZproxy licences will be required, the precise number depending on the required service capacity. For conventional use, purchase is simple and an institution would pay \$495 per server. Operation as a centralised service on behalf of other

institutions may be less familiar territory for OCLC and it may take some time to negotiate a price scale for various levels of service capacity (see <http://www.usefulutilities.com/purchase/>).

- e) *Opportunity cost*. A further cost, again difficult to quantify, is the lost opportunity for further UK federation development. The effort spent developing and improving the proxy solution comes directly at the expense of the staffing resource available for core federation activity.
- f) *Service maintenance*. Any centralised service requires a support service set up to maintain the software and configuration information and to answer user queries regarding the service.
- g) *Time*. All of these items will cost time as well as effort. Time is in short supply given that the service is required by Midsummer.

5 Risks

The risks at various levels are considered below. They are also summarised in Table 1.

5.1 Deployment

The risk of failing to deploy a functional central facility on time is largely dependent on a number of external factors. These include the following elements, which must be in place on time:

- equipment procurement, negotiation and delivery;
- establishing equipment in colocations;
- negotiating EZproxy licence with OCLC;
- negotiation with service providers to support the proxy;
- promotion of the service to institutions and take-up;
- technical development of EZproxy Central;
- development of procedures and guidance notes for institutions, service providers, and the proxy operator.

Only the last two of these dependencies is within the immediate control of the developers.

5.2 Take-up

There is no guarantee that the community will adopt the solution. For the Librarian, a centralised EZproxy facility is not a substitute for a bespoke institutional proxy, which would be tailored according to the presentation style of the institution and the resources actually available to its users. A 'one size fits all' solution, where users are presented with inappropriate service choices, may be seen as unacceptable.

The solution does not support user personalisation, which for some resources may be regarded as a key element of the service. Where an institution regards even a single resource of this type as critical, this may be sufficient cause for it to continue using OpenAthens. In this case, the institution would have little interest in using EZproxy Central.

The look and feel of the proxy is likely to be contentious, with as many opinions as there are participating institutions; inevitably, there will be a lobby for institutional customisation. Additional effort spent on incremental improvements to the proxy would soon yield diminishing results.

The tight timetable does not allow institutions much time to consider their deployment options or to hold user trials. In the absence of hard information on the effectiveness and usability of the service, the conservative option would be to pay for an OpenAthens subscription.

Unlike Shibboleth, or indeed Athens, the solution is not transparent. The user cannot visit the service resource first, and present credentials later; the user must always login to the proxy first, and link to the

required service from there. This is routine behaviour in environments where users access resources through an institutional portal, but is less suitable where users are accustomed to service-first access.

The cooperation of service providers cannot be taken for granted though the virtualised proxy solution may be the least alarming to them. Service providers must be satisfied that their commercial interests will not be compromised and that the proxy's assertions about the origins of service requests can be trusted.

5.3 Exit strategy

If the solution did prove successful, with wide adoption by institutions and acceptance by service providers, an unintended consequence could be the removal of incentives to adopt UK federation standards; An IP-based solution, while being functionally limited, has the appeal of familiarity and low maintenance costs. A successful service may also lead to pressure from both institutions and service providers to make the facility a permanent part of the access infrastructure. This would entail an ongoing cost against resources that could otherwise be used to develop the UK federation.

An alternative outcome may be an encouragement to institutions to deploy EZproxy services locally and a consequent removal of community pressure on service providers to support access by Shibboleth. At the least, there would be some dilution of the message to implement Shibboleth.

Table 1— Summary of risks

Availability within timetable	This is a significant infrastructure and support investment to be in place in a very limited time frame against a hard deadline. All of the technical options have significant drawbacks.
Unquantifiable size	There is no reliable information on which to base user or resource provider uptake, nor to quantify the load the service would need to accommodate. If sized too small the user experience would be poor and the service would attract significant flak from institutions. If sized too large the per-institution cost would be high and raise questions of value for money.
Alternative solution	OpenAthens offers a familiar alternative at a price that some institutions may consider affordable.
Impact on growth in use of federation core technology	If successful, the proxy service removes incentives from both institutions and service providers to move to using federation core technology.
Usability limitations impact on take-up	Some of the usability issues may be regarded as critical limitations and may block take-up: lack of user personalisation; 'proxy-first' visit to services; presentation of inappropriate service choices; proxy look and feel.
Opportunity cost	Unless particularly successful, the development to service could well consume time and money that could have been more effectively used for core federation development and promotion.
Acceptability to resource providers	Some providers may require persuasion to become convinced of the security of the service. A different issue is that some services require user accountability and do not accept IP address checking in any form.
Exit strategy	If successful, closing the proxy service is likely to meet with significant resistance; continuing with it has cost implications and consequences for the migration of institutions and service providers to federation core technology.

6 Interim conclusions

1. Provision of an EZproxy Central service carries a number of substantial risks and has the distinct possibility of being an expensive failure. On present evidence, the benefits appear speculative and are outweighed by the risks which are all too apparent.
2. If it is decided to investigate the scheme more fully, further evidence could be collected:
 - consult colleagues involved in the original investigation;
 - test the assumptions within the technical options, and document the development steps in detail;
 - undertake analysis of service usage, critical service requirements (such as personalisation);
 - investigate institutional views on take-up, in the light of usability concerns.
3. If it is decided that a central facility is desirable, then given the extent of the work necessary a more realistic plan would be to schedule the development of the EZproxy Central solution for use in academic year 2009-2010.
4. An alternative solution may be to offer support to institutions willing to implement their own EZproxy service. Such support may include negotiation of a bulk purchase of EZproxy from OCLC as well as advice on technical installation and ongoing help-desk support.