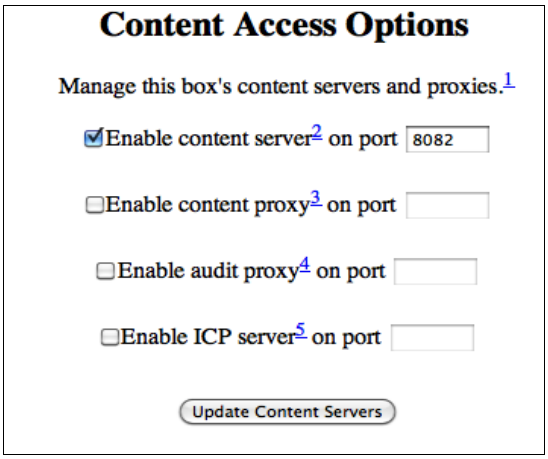
Lots of Copies Keep Stuff Safe (LOCKSS) and Innovative Interfaces have made content preserved on LOCKSS Boxes available through OPACs that use the WebBridge LR link resolver. This guide describes the steps required to customize your LOCKSS Box and WebBridge LR instance to integrate seamlessly the content preserved on your LOCKSS Box into the user's online journal searches.

Customizing Your LOCKSS Box

A LOCKSS Box has an optional Content Server that makes preserved content available from a web browser through OpenURL queries. The SFX LOCKSS target uses OpenURL queries to request content from your local LOCKSS Box. The Content Server is not enabled by default. To make content in your LOCKSS Box available through SFX, use its configuration screens to enable and configure this feature.

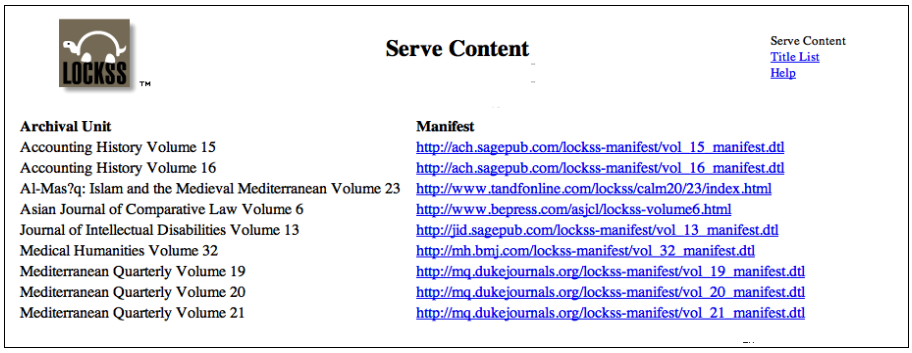
Enabling the Content Server

From the main administration screen of your LOCKSS Box, select **Content Access Options**. Then select **Content Server Options**. This will take you to a screen for managing the LOCKSS Box's content servers and proxies. Select **Enable content server on port** and enter the port number to use for serving preserved content from a web browser. Port numbers below 1024 require the application to run as root and are therefore not available. It is recommended that you choose a port that is close to 8081, the default port used by the administration GUI of of your LOCKSS Box. Port 8082 is used in this example.



Finally, select **Update Content Servers** to finalize this change. The operating system on your LOCKSS Box host may require opening non-standard ports to enable outside access. Consult with the person who administers the host where your LOCKSS Box is running to determine which standard ports are available.

To test your Content Server configuration, use your browser to contact the Content Server of your LOCKSS Box which you set up above. By default, the Content Server displays a list of Archival Units (AUs) that are configured. In this example, if your LOCKSS Box administration GUI is accessed at http://lockss.xyz.edu:8081/, your Content Server is at http://lockss.xyz.edu:8082/ServeContent.

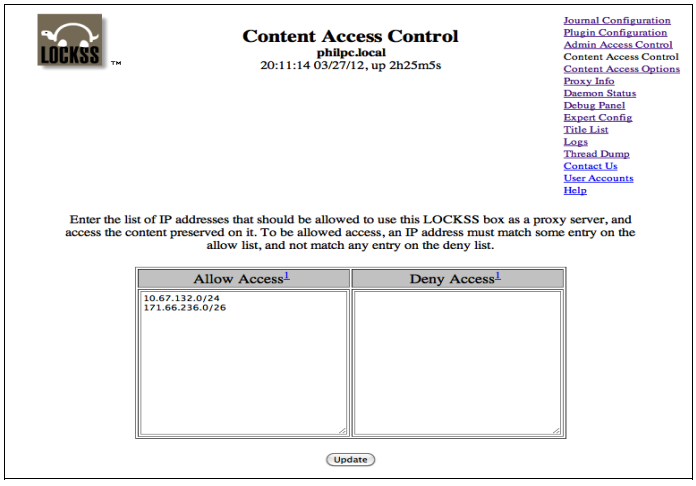


You must ensure that the packet filters (i.e. firewall) configured on your LOCKSS Box allow access to the port you chose for your Content Server. The default packet filter is 8082. Enabling the port for access outside of the network requires administrative privileges, and depends on the operating system running your LOCKSS Box. Ask the systems administrator to do this.

Configuring Content Server Access Control

Now that you have enabled the Content Server, you'll need to give your user community access to preserved content. To do this, return to the main administration page of your LOCKSS Box and select **Content Access Control**. You will see two lists side-by-side. The "Allow Access" list specifies which IP addresses or ranges can access preserved content through the Content Server. The **Deny Access** list allows you to deny access to specific IP addresses or ranges, typically ones within a wider range specified by the **Allow Access** list.

The default **Allow Access** list includes two address ranges: the initial network address that was specified when your LOCKSS Box was installed, and those by the LOCKSS organization to provide technical support.



Consult your contracts with publishers whose content is preserved on your LOCKSS Box to determine what IP address ranges have access to them. You should configure the **Allow Access** list to include the same IP address ranges that you supplied to publishers.

Note: The simple configuration that is supported by your LOCKSS Box assumes that all publishers support the same IP address ranges.

For example, XYZZY University provides the following IP address ranges to its e-pub vendors:

IP Address	CIDR Prefix	Netmask	Use
10.12.0.0	/16	255.255.0.0	Dorms
10.64.0.0	/16	255.255.0.0	Main Campus
10.65.0.0	/16	255.255.0.0	Medical Center

Your institution's contracts with publishers may exclude certain campus functions. Only those address ranges that are allowed should be included in the **Allow Access** list. Here is what XYZZY University would add to its list:

10.12.0.0/16

10.64.0.0/16

10.65.0.0/16

Once you are done configuring the access control for your LOCKSS Box, select **Update** to finalize your configuration. Test your configuration by accessing the Content Server from various allowed and denied hosts within your institution.

Using Apache as a Front End to your LOCKSS Content Server

The configuration just described gives end users direct access to your LOCKSS Box's Content Server through the port you specified. Some IT organizations limit the use of “non-standard” ports for web services because it requires opening additional ports in their institution's firewalls. These organizations may also prefer to administer access using the same web server access controls that they use throughout their institution. If this does not apply to your institution, you can skip this section.

One way to accomplish this is to direct users to a web server such as Apache that runs on the standard port 80, and configure the Apache server to act as a proxy. It relays requests to the LOCKSS Box Content Server and returns responses from the LOCKSS Box to the user. This allows the presence of the LOCKSS Box to be hidden, and enables the IT staff to use existing access control configurations on the Apache server to limit access to preserved content, including user authentication using HTTPS and SSL certificates.

Load balancing could also be done across several LOCKSS Boxes in a similar way. The Apache instance can be one that is shared by many applications within your institution, or it can be a custom instance that is installed on the same host as the LOCKSS Box. The simplest configuration uses the standard 'mod_proxy' Apache module to forward content server traffic to and from the LOCKSS Box. Here is a typical 'mod_proxy' configuration for an Apache server virtual host at port 80 on the same machine as the LOCKSS Box:

```
<VirtualHost *:80>
ServerAdmin lockss-admin@xyzzy.edu
DocumentRoot /var/www/html
ServerName lockss.xyzzy.edu
<IfModule mod_proxy.c>
ProxyRequests Off
ProxyVia On
<Proxy lockss.xyzzy.edu/*>
AddDefaultCharset off
Order deny,allow
Allow from all
</Proxy>
ProxyPassMatch ^/((ServeContentlimages).*)$ http://localhost:8082/$1
</IfModule>
ErrorLog logs/error_log
CustomLog logs/access_log common
</VirtualHost>
```

If you use this technique, you should omit adding IP addresses to the **Content Access Control** of your LOCKSS Boxes because it will only be accessed locally by the Apache server. If the Apache server is on a different subnet, you should add only the IP address of that host to the Content Access Control of your LOCKSS Box. You should also add the following special configuration parameter to **Expert Config** screen of your LOCKSS Box.

```
org.lockss.serveContent.absoluteLinks=false
```

This causes Content Server to rewrite links in pages it serves relative to the LOCKSS Box address, which simplifies the 'mod_proxy' configuration.

Customizing Your WebBridge LR Interface

Once you have configured your LOCKSS Box, the next step is to make its contents available through WebBridge LR.

Getting Coverage Information from the LOCKSS Box

To find out what journal content has been preserved on a LOCKSS Box, navigate to the Administration page (e.g. http://lockss.xyz.edu/) and click the **Title List** link in the top right corner. The Title List will be generated as a spreadsheet and can show various types of content. Be sure to click the *Collected* button in the **Show** field, the *Title Ranges* in the **Format** field, and the *TSV* button in the **Output** field. Then click the **List Titles** button at the bottom of the page.

This will generate a Tab Separated Value file which can be opened with any spreadsheet software (i.e. Excel) and will list the Journal Titles preserved on the LOCKSS Box as well as the year ranges of each Title. You will need to prepare the file for the coverage data load process to your Innovative System in the same manner you would prepare any other data file for loading.

URL Syntax for Linking to Your LOCKSS Box

The LOCKSS Box will accept any URL formatted to the OpenURL standard. The base URL is:

http://<yourLOCKSSbox>/ServeContent?

And a typical link URL for WebBridgeLR might look like this:

http://<yourLOCKSSbox>/ServeContent?
issn=#@ISSN#&volume=#@VOLUME#&issue=#@ISSUE#&spage=#@SPAGE#

Note: The LOCKSS Boxes support both hyphenated and non-hyphenated ISSN formats, support both OpenURL 0.1 and 1.0 syntax, and follow ISO formatting for dates with month and day being optional.

Successfully Linking to LOCKSS Content

Since the LOCKSS Boxes use OpenURL format, any metadata can be used to link to content: **ISSN**, **Volume**, **Issue**, and **Start Page** should be the bare minimum metadata necessary to reach a specific article.

Note: Other metadata which may assist in reaching a specific article might include **Article Title** and **Article Number**, however these metadata are not as widely supported as the **Start Page**.

If any metadata is requested that isn't available, the LOCKSS Box will take the user to a top-level page which lists all of the preserved volumes for a particular Journal Title so that they can navigate by hand from there.