LOCKSS User Interface (UI)

http://mycache.edu:8081

Page Name	Explanation Page	Function
Journal Configuration	YES	Add AUs to the cache for ingestion. Remove an AU from preservation activity if needed.
Admin Access Control	YES	List IP addresses that are allowed to access the UI.
Content Access Control	YES	List IP addresses that are allowed to access content held on this cache.
Content Access Options	YES	Configure how content is collected from publishers and served to users.
Proxy Info	YES	Obtain proxy configuration information for browsers and other proxies, to inform them which URLs should be proxied through this LOCKSS box.
Daemon Status	YES	Detailed information about the Archival Units (AU) and voting and polling and configuration of this cache are accessed via the Daemon Status pages. See the Daemon Status Explanation Pages.
Debug Panel	YES	Force the daemon to take certain immediate actions. This is usually done for testing and troubleshooting and is not part of normal everyday operations.
Expert Config	No	Set parameters not available in the UI; settings survive a daemon restart.
Logs	No	Links to view the Daemon log files.
Thread Dump	No	Advanced Class
Contact US	No	Send Email to MetaArchive Support. (This is configurable in the title database.)
Help	No	Hyperlink to Lockss Help Pages

JOURNAL CONFIGURATION

http://mycache.mysite.edu:8081/BatchAuConfig

Selection	Data Type	Meaning	Inputs/Actions	Special Comment
Add Titles	Hyperlink	Select AUs to be preserved on this cache.	Check boxes and submit buttons.	After a title is selected for preservation, it will no longer be listed as an item which can be preserved.
Remove Titles	Hyperlink	Select AUs this cache will no longer preserve. The title will be listed only if the cache is already preserving it.	Check boxes and submit buttons.	The daemon does not and can not remove any of the AU's preserved data that is stored on the disks of the cache.
Deactivate Titles	Hyperlink	Select AUs for which this cache will temporarily suspend preservation activiites.	Check boxes and submit buttons.	The AU will be listed only if the cache is already preserving it.
Reactivate Titles	Hyperlink	Select AUs for which this cache will continue to resume preservation activiites.	Check boxes and submit buttons.	The AU will be listed if the cache has deactivated it.
Backup	Hyperlink	Create a "BatchAUconfig" file with this caches journal configuration settings. This is NOT a backup of the data on the cache. Only active AUs are backed up. To keep backup info on any AUs that are deactivated, temporarily reactivate them before making the backup.	You will be asked to save the file to the disk of your PC. The backup is a single .zip compressed file.	Our caches are backed up on a monthly basis to the central servers.
Restore	Hyperlink	Restore the cache journal configuration if it is has been lost or corrupted.	You will be asked for the location of the file on the disk of your PC.	A cache can also be 'cloned' using this method.
Manual Add/Edit	Hyperlink	Edit AUs on the cache, and can work around the title database.		Not used in production.

Admin Access Control

http://mycache.mysite.edu:8081/AdminIpAccess

Selection	Data Type	Meaning	Inputs/Actions	Special Comment
Page Explanation		As a security feature the LOCKSS UI will only respond to known IP addresses. This page is used to enter the IP addresses that are allowed to interact with the the UI. By default, no IP addresses are allowed except those that are initially configured as allowed when configuring the cache.	Standard IP address notation. Explicit address: aa.bb.cc.dd or subnet: aa.bb.cc.0/#	Only a user from a known IP address can add/delete new IP addresses. User/Pass authentication is then needed.
Allow Access	Text List	Enter the list of IP addresses that should be allowed to access the administrative user interface on this LOCKSS box. To be allowed access, an IP address must match some entry on the allow list, and not match any entry on the deny list.	Addresses are added here, more than one can be added.	You must click on update and wait for a 'update success' message to actually modify the list. (These can be set in the title database as well.)
Deny Access	Text List	Enter the list of IP addresses that are not allowed to access the administrative user interface on this LOCKSS box. To be allowed access, an IP address must match some entry on the allow list, and not match any entry on the deny list.		This parameter is rarely used, but is available should a situation arise where it is needed.

Content Access Control

http://mycache.mysite.edu:8081/ProxyIpAccess

Selection	Data Type	Meaning	Inputs/Actions	Special Comment
Page Explanation		As a security feature the LOCKSS cache will ONLY allow content to be served out to known IP addresses. This page is used to enter the IP addresses that are allowed to retrieve content that is contained on the cache. By default, no IP addresses are allowed.	Standard IP address notation. Explicit address: aa.bb.cc.dd or subnet: aa.bb.cc.0/#	In MetaArchive, a "dark" archive, an address is only added when a retrieval is needed. After retrieval the address is deleted.
Allow Access	Text List	Enter the list of IP addresses that should be allowed to use this LOCKSS box as a proxy server, and access the content preserved on it. To be allowed access, an IP address must match some entry on the allow list, and not match any entry on the deny list.	Addresses are added here, more than one can be added.	You must click on update and wait for a 'update success' message to actually modify the list.
Deny Access	Text List	Enter the list of IP addresses that are not allowed to use this LOCKSS box as a proxy server, and access the content preserved on it. To be allowed access, an IP address must match some entry on the allow list, and not match any entry on the deny list.		This parameter is rarely used, but is available should a situation arise where it is needed.

Content Access Options

http://mycache.mysite.edu:8081/ProxyAndContent

Selection	Data Type	Meaning	Inputs/Actions	Special Comment
Page Explanation		This page is used when special configuration of the cache is needed for the cache how to access content that it is ingesting and/or how to allow access to content which it has ingested/preserved.		
Content Server Options	Hyperlink	Settings for managing how this cache serves content. (3 sub pages are needed for the 3 types)	Select the ports that active servers are listening on.	Remember that firewalls and the end user need to know these port numbers.
Proxy Client Options	Hyperlink	Configure the LOCKSS crawler to access the net through a proxy server.	someserver.somelibrary .edu 8080	URL and Port: Leave off the http:// Check Proxy crawls.
Export Content	Hyperlink	Export preserved content as Zip, WARC, etc.	Select the AU and the export type.	After entering the details click the Create Export File button. When the archive is ready a link to it appears on the page.

Proxy Info

http://mycache.mysite.edu:8081//info/ProxyInfo

Selection	Data Type	Meaning	Inputs/Actions	Special Comment
Page Explanation		LOCKSS caches are inherently capable of acting as PROXY servers. These selections are used to help create proxy configuration files, so an end user can retrieve content from the cache, even though they are using the URL of the original source. There are a handful of industry standard proxy server formats supported here.	In all cases select the radio button and click the SUBMIT button. The data is returned to the browser.	MA is a dark archive. Opening up content is only used for testing, and retrieval of lost data. The functions on this page are rarely used in production.
EZ Proxy	Radio Button	http://www.oclc.org/ezproxy/		Extensively used by libraries to give access from outside the library's computer network to restricted-access websites that authenticate users by IP address.
Squid	Radio Button	http://www.squid-cache.org/		
PAC File	Radio Button	http://en.wikipedia.org/wiki/Proxy_auto-config	MA typically uses this technology when testing proxy out capability.	Automatic proxy configuration for browsers. Place the contents of this file on a server for your users to configure their browsers.

Debug Panel

http://mycache.mysite.edu:8081/DebugPanel

Selection	Data Type	Meaning	Inputs/Actions	Special Comment
Page Explanation		This page allows the user to perform direct actions on the cache that take place immediatley, but are normally run by the cache on a scheduled basis.	Buttons are selected to take the action	It is rare that these items are used in production. But they are regularly used in the test network.
Reload Config	Button	This will cause the cache to reload the Title Database. Caches reload their title database on a regular basis (which is defined by org.lockss.config.reloadInterval). This is how the cache can learn about new AUs available for ingest and daemon parameter changes.		This is handy to force a daemon to reload the title database immediately after the the title database has changed.
Mail Backup File	Button	This mails the BatchAuConfig which is the same as the Backup on the Journal Config page.	Can be mailed on a regular basis.	MA backs up the config files of all production caches monthly.
Hash Cus	Hyperlink	Advanced Class		
Throw IO Exception	Button	Advanced Class		
Start V3 Poll	Button	Request that the integrity of the selected AU be checked against the cache's peers.	Choose the AU from the pick list to take this action	In some cases you will need to Over ride the 'Rate Limiter' and click the button a second time.
Start Crawl	Button	Start a crawl of the AU	Choose the AU from the pick list to take this action	In some cases you will need to override the 'Rate Limiter' and click the button a second time. Note - This CAN NOT be used to override a crawl window**.

Daemon Status Pages

http://mycache.edu:8081/DaemonStatus

Page Name	Useful for Troubleshooting?	Special Comment
AU Ids	A little	Lists all preserved AUs on the cache, with links to each AUs detail page. Maps AU names/titles to the identifying string (AUId) made up of the plugin id, and the AU's parameter values. The AUId is a unique identifier by which LOCKSS daemons recognize AUs.
All Title AUIDs	A little	Lists all AUs known to the cache, the actively preserved AUs and inactive AUs. This is not to be confused with the AU Ids list, or the AU Ids in the title database. AU Ids listed here and not listed in the AU Ids list are either deactivated or, listed in the title database but not preserved by this cache.
Archival Units	YES	Links to details of the actively preserved AUs
Comm Channels	A little	Useful to check status of SSL over LCAP. Useful to quantify non-crawling related network bandwidth consumed by the cache.
Comm Peer Data	A little	Useful if see LCAP issues
Comm Statistics	NO	Should be ignored in MA
Configuration	Should never be an issue in production	LOCKSS settings from the title DB and the caches configuration file. While there should never be any surprises in a production network, this is useful to check whether the expectations are reality.
Crawl Status	YES	Verify success/failure/progress of content ingest.
Hash Queue	YES	Analyze time spent on fixity checks
Overview	Quick Status Overview	Default Page when go to Daemon Status
Peer Identities	A little	These are fixed in MA
Platform Configuration	Should never be an issue in production	Handful of LOCKSS settings specific to this cache
Polls	YES	Status of Polls initiated by this cache
Publisher Plugins	Should never be an issue in production	Plugins this cache knows about
Repositories	A little	Map of file system directory to AU location
Repository Space	A little	Check if disk space issues
Scheduler Queue	Advanced Class	
Timer Queue	Advanced Class	
Votes	YES	Status of Votes in which this cache was asked to participate

ARCHIVAL UNITS

http://mycache.mysite.edu:8081/DaemonStatus?table=ArchivalUnitStatusTable

Field Name	Data Type	Meaning	Example	Special Comment	Purpose
Volume	Text Description	AU name and parameter (If Any)	Folger Digital Image Masters: from=19451	The name is a hyperlink to details of the AU.	Alphabetical listing of AUs on this cache.
Content Size	Positive Integer	Size in bytes always. Not KB/MB/GB.	6,201,385,391		
Disk Usage (MB)	Positive Integer	How much disk space is used in MB.	5922.3	Disk usage can be greater than the content size. This is because it also includes the space used for earlier revisions of an AU.	
Peers	Hyperlink	Shows the peers that will repair this AU.			
Recent Polls	Hyperlink	Polling history on this AU		The maximum time between polls is set in the Title Database.	
Status	Percentage	Agreement percentage. In general 100%. (Flags may follow status: C means the AU is complete, D means that the AU is no longer available from the publisher, NS means the AU has no files containing substantial content.)	100.00% Agreement	Agreement less than 100% is OK for AUs that have changing data.	Sort on this column to find troubled AUs. Agreement < 90%
Last Poll	Time/Date Stamp	Most recent poll requested for this specific AU.	23:48:57 09/16/10	Should not be > 8 weeks.	
Last Crawl Start	Time/Date Stamp		09:52:56 09/23/10		
Last Crawl Result	Status (Text)	If the last crawl succeeded, or an error status.	Successful | Fetch error | Can't fetch permission page	Fetch errors are not unusual due to the nature of Internet Web crawls. A permission page error needs to be looked	Sort on this column to find troubled crawls.
Last Successful Crawl	Time/Date Stamp		10:18:17 09/23/10		

Guide to the LOCKSS User Interface

Crawl Status

http://mycache.mysite.edu:8081/DaemonStatus?table=crawl_status_table

Field Name	Data Type	Meaning	Example	Special Comment	Importance
Journal Volume	Text Description	AU name and parameter (If Any)	Folger Digital Image Masters: from=19451	The name is a hyperlink to details of the AU.	Alphabetical listing of AUs on this cache.
Crawl Type	Text Description	New Content or Repair	New Content	Repairs are rare	
Start Time	Time/Date Stamp	When the crawl was started. History of more than just the last crawl is usually shown.	17:37:55 09/16/10	You can see a problem one day cleared the next.	
Duration	Elapsed Time	How long it took this cache to crawl this AU	5h9m58s	Can be ms to days	
Status	Hyperlink	Very brief status of the crawl		Great for specific errors when crawl fails	Can help locate problems in the AU source
Bytes Fetched	Positive Integer	Number of content bytes collected from server during crawl. Does not include HTTP headers or other network overhead.	16,338,119	Size in bytes always. Not KB/MB/GB.	
Pages Fetched	Positive Integer	Number of pages successfully fetched from server	885		Pages are fetched and then run through the plugin rules.
Pages Parsed	Positive Integer	Number of (html, etc.) pages were run through the plugin rules and scanned for URLs	884	LOCKSS parses a handfull of formats, HTML, CSS, ... but not XML, images, word documents, …	Plugin rules determine content that is PRESERVED.
Pages Pending	Positive Integer	Number of pages waiting to be fetched	Successful		
Pages Excluded	Positive Integer	Number of pages that didn't match the crawl rules	0	LOCKSS parses a handfull of formats, HTML, CSS, ... but not XML, images, word documents, …	
Not Modified	Positive Integer	Number of pages for which we already had current content	1		
Errors	Positive Integer	Number of pages that could not be fetched	0	Hyperlink to errors	Sort on this column to find troubled crawls.
Mime Types	Positive Integer	Number of different content types	2	Hyperlink to types (i.e.): application/pdf 1370 text/html 1372 text/xml 5	

Guide to the LOCKSS User Interface

Polls

http://mycache.mysite.edu:8081/DaemonStatus?table=V3PollerTable

Field Name	Data Type	Meaning	Example	Special Comment	Importance
Volume	Text Description	AU name and parameter (If Any)	Folger Digital Image Masters: from=19451	The name is a hyperlink to details of the AU.	Alphabetical listing of AUs on this cache.
Participants	Positive Integer	Number of caches that participated in the poll.	4	Always 3 or more to reach quorum.	Proves cache is communicating with peers.
Status	Text Description	Status of a poll if not complete, or the results of the poll after completion. Ususally complete.	Complete	Sorting on this column will show AUs that are not reaching quorum.	No Quorums are a regular occurance, but should not go on for more than a few days.
URLs Tallied	Positive Integer	Total number of URLs examined so far in this poll	2745		
Hash Errors	Hyperlink	Errors encountered while hashing content		Almost always non-existent	
Repairs	Positive Integer	Completed repairs	0		
Agreement	Positive Integer	Percentage agreement	100%	Below ~90% agreement signifies some kind of issue	Sort on this column to find troubled Aus. Agreement < 100%.
Start	Time/Date Stamp	When the poll will or did start.	07:33:21 10/01/10		
Deadline	Time/Date Stamp	When the poll is scheduled to end	17:35:23 10/09/10	The duration is not hard and fast. But polls will be terminated if lasting longer.	
Poll ID	Alphanumeric	Random string	CLTRGSZnKJ	Hyperlink to details of the Poll	Link shows which caches were in the last poll.

Guide to the LOCKSS User Interface

Votes

http://mycache.mysite.edu:8081/DaemonStatus?table=V3VoterTable

Field Name	Data Type	Meaning	Example	Special Comment	Importance
Volume	Text Description	AU name and parameter (If Any)	Folger Digital Image Masters: from=19451	The name is a hyperlink to details of the AU.	Alphabetical listing of AUs on this cache.
Caller	IP Address	IP address of the cache that called for the poll.	TCP:[128.173.125.42]: 9729	IP addresses are what is in the title database.	Proves cache is communicating with peers.
Status	Text Description	Status of voting in the poll.	Complete | Hashing | Expired w/o Voting	Usually complete or No Time Available	There are phases: Accepted, hashing, voted, complete.
Start	Time/Date Stamp	When the Poll was started by the caller.	17:48:04 10/03/10		
Deadline	Time/Date Stamp	When the voting must be done.	18:30:26 10/03/10	Not voting by the deadline leads to an "Expired w/o Voting" or Error Status	Too many Expired w/o Voting is a problem.
Poll ID	Alphanumeric Hyperlink	Random string	CLTRGSZnKJ	Clicking the hyperlink reports details of the Poll and which cache called for the Poll	

Guide to the LOCKSS User Interface

Hash Queue

http://mycache.mysite.edu:8081/DaemonStatus?table=HashQ

Pending requests are first in table, in the order they will be executed. Completed requests follow, in reverse completion order (most recent first).

Field Name	Data Type	Meaning	Example	Special Comment	Importance
Req	Sequential Integers	Simple counter	406		Order in which requests were made
State	Text Description	Where in the process is the hashing calculation, is done, or did it timeout, etc.	Done | Timeout		
Volume	Text Description	AU name and parameter (If Any)	Virginia Libraries: volume=52	NOT a hyperlink here	
Cached Url Set	Text Description	Describes if the entire AU is being hashed (AUCUSS), or a subset of the AU is being hashed (SNCUSS) as part of the repair process.	[AUCUSS] | [SNCUSS]	The entire AU, or a piece of the AU if a repair is in process.	
Type	Text Description	Hashing can take place for the obvious reason, to participate in polling and voting, but it also takes place to help create estimates for how long a vote should take.	B(2) | E(2)		Hashing to create an estimate for the voting process is part of cache overhead.
Deadline	Time/Date Stamp	Time by which hashing must be completed	17:18:04 04/16/10		Vote ends in error if not completed in time.
Estimated	Time Duration	How long it should take to hash the AU and produce a fixity value (checksum).	4m41s		
Used	Time Duration	How long it did take to hash the AU and produce a fixity value.	2m36s	Displayed In Red if > estimate	
Bytes Hashed	Positive Integer	Bytes Hashed	10,292,191,034	In Bytes	

Peer Identities

http://mycache.mysite.edu:8081/DaemonStatus?table=Identities

Sort on Messages http://mycache.mysite.edu:8081/DaemonStatus?sort=DorigTot&table=Identities

Field Name	Data Type	Meaning	Example	Special Comment	Importance
Peer	IP Address	Identifies the peer cache	TCP:[128.173.125.42]:9729	Peers listed in simple TCP format do not participate in polls.	
Last Message	Time/Date Stamp	Last time a message was received from this peer.	22:27:46 04/16/10	Records will go back for the history of the cache.	Records will show peers that have been in contact with this cache. It does not matter when the contact was.
Message Type	Text Description	Last message type that was sent from this peer.	Poll (10) | PollAck (11) | Nominate (13) | VoteRequest (14) | Vote (15) | EvaluationReceipt (18)	Several types exist	
Messages	Positive Integer	Total number of messages received from this peer.	6009	Records will go back for the history of the cache.	Sorting on this column is very helpful to see where there is lots of activity with a peer.
Last Poll	Time/Date Stamp	Last time that the local peer participated in a poll with this peer.	14:16:37 08/22/10		
Last Vote	Time/Date Stamp	Last time that this peer participated as a voter in a poll called by the local peer.	06:47:52 08/25/10		
Last Invitation	Time/Date Stamp	Last time this peer was invited into a poll called by the local peer.	01:29:40 09/01/10		
Invitations	Positive Integer	Total number of invitations sent to this peer by the local peer.	3522	Records will go back for the history of the cache.	
Polls Called	Positive Integer	Total number of polls called by this peer in which the local peer voted.	783	Records will go back for the history of the cache.	Polls & votes give an idea of how busy this cache is working with this peer.
Votes Cast	Positive Integer	Total number of polls called by the local peer in which this peer voted.	921	Records will go back for the history of the cache.	
Polls Rejected	Positive Integer	Total number of poll requests rejected by this peer	42	Rejects for not having the AU is to be expected.	
NAK Reason	Text Description	Reason for most recent poll request rejection, if any.	Too Busy | No AU		
Groups	Text Description	For now there is only the MA group	metaarchive	A cache can belong to more than one group.	