

Infrastructure - Task #2072

Story # 2067 (Closed): Refactor the index task generator and processor to improve scalability and performance

Update the index task generator to listen to the hzSystemMetadata map for updated entries

2011-11-29 14:15 - Dave Vieglais

Status:	Closed	Start date:	2011-11-29
Priority:	Normal	Due date:	
Assignee:	Skye Roseboom	% Done:	100%
Category:	d1_indexer	Estimated time:	0.00 hour
Target version:	Sprint-2012.05-Block.1.3	Story Points:	
Milestone:	CCI-1.0.0		
Product Version:	*		

Description

The index task generator currently iterates through the entire map of system metadata entries, which is very inefficient especially as the number of entries increases.

The task generator should implement a listener that creates a new indexing task in response to creation or update of system metadata.

An example listener is available in the hzPeek tool in subversion at:

https://repository.dataone.org/software/cicore/trunk/cn/d1_cn_tools/d1_hzpeek/src/main/java/org/dataone/cn/utility/hzpeek/

Note that a single sys meta entry may be updated several times in succession. The index task queue should be examined for existing entries for the same PID, and those should be purged if present, to be replaced by the most recent entry.

History

#1 - 2011-12-01 20:21 - Skye Roseboom

- Status changed from New to In Progress

Updated generator to remove existing tasks that have the same systemMetadata PID and have task status of 'new'.

If tasks exists with same PID and process is 'in-process', 'failed', or 'completed' - no duplicate is removed and a new task is generated.

#2 - 2011-12-07 21:15 - Skye Roseboom

- Status changed from In Progress to Closed

Created an EntryListener implementation for the systemMetadata hazelcast map. This listener is wrapped in a apache commons daemon for jsvc support.