Infrastructure - Bug #4204

cn.archive throws errors on DATA objects

2013-12-17 05:15 - Rob Nahf

Status:	Closed		**Start date:**	
Priority:	Normal		**Due date:**	
Assignee:	Chris Jones		**% Done:**	100%
Category:	Metacat		**Estimated time:**	0.00 hour
Target version:				
Milestone:	CCI-1.2		**Story Points:**	
Product Version:	*			

Description

When attempting to reflect an mn.archive event during synchronization using cn.archive (TransferObjectTask.updateSystemMetadata()), a NotFound exception is thrown by the call to metacat, at D1NodeService.archive(). for example:

```
135538  [ INFO] 2013-12-13 21:18:43,476 (TransferObjectTask:write:430)
Task-urn:node:mnDemo1-syncTesting_Obs_n_Archvd:2013347131323615 Pid Exists. Must be an Update
135539  [ INFO] 2013-12-13 21:18:43,477 (TransferObjectTask:write:462)
Task-urn:node:mnDemo1-syncTesting_Obs_n_Archvd:2013347131323615 Update sysMeta because checksum is same
135540  [ INFO] 2013-12-13 21:18:43,479 (TransferObjectTask:updateSystemMetadata:687)
Task-urn:node:mnDemo1-syncTesting_Obs_n_Archvd:2013347131323615 Update ObsoletedBy
135541  [ INFO] 2013-12-13 21:18:43,945 (TransferObjectTask:updateSystemMetadata:692)
Task-urn:node:mnDemo1-syncTesting_Obs_n_Archvd:2013347131323615 Updated ObsoletedBy
135542  [ INFO] 2013-12-13 21:18:43,945 (TransferObjectTask:updateSystemMetadata:697)
Task-urn:node:mnDemo1-syncTesting_Obs_n_Archvd:2013347131323615 Update Archived
135543  [ERROR] 2013-12-13 21:18:44,182 (TransferObjectTask:write:488)
Task-urn:node:mnDemo1-syncTesting_Obs_n_Archvd:2013347131323615
135544  <?xml version="1.0" encoding="UTF-8"?>
135545
135546     The object with the provided identifier was not found.
135547
135548

135549  [DEBUG] 2013-12-13 21:18:44,950 (TransferObjectTask:call:205)
Task-urn:node:mnDemo1-syncTesting_Obs_n_Archvd:2013347131323615 Unlocked task
```

The culprit is the check in D1NodeService.archive() that the pid exists in the IdentifierManager:

```
// check for the existing identifier
try {
localId = IdentifierManager.getInstance().getLocalId(pid.getValue());
} catch (McdbDocNotFoundException e) {
throw new NotFound("1340", "The object with the provided " + "identifier was not found.");
}
```

The reason the error happens is that DATA objects are not created on the CNs, only the system metadata is added via registerSystemMetadata, and only to the Hazelcast sysmeta map:
TransferObjectTask.createObject() calls nodeComm.getCnCore.registerSystemMetadata(), which is handled by CNReadController in d1_cn_rest_proxy, and ultimately ends up at CNodeService.registerSystemMetadata(), which only updates the hazelcast sysmeta map.

This results in an inconsistency between the extent of pids in the hz sysmeta map (which include DATA, METADATA, and RESOURCE formats), and the metacat identifier table on the CNs (only METADATA and RESOURCE format pids).   There seems to be also an inconsistency in what the locus of control is for systemMetadata when hz is involved (metacat as CN) - archive checks with, and acts on the database tables, while other methods (setObsoletedBy, for example) only act on the hz systemMetadata map.

(As an aside, It's also unclear to me how changes made to the hz systemMetadata map are propagated to the underlying data store.)

Are there other CN methods that update systemMetadata that similarly bypass the hazelcast map that might have similar issues?

Related issues:			
Blocks Infrastructure - Story #4193: Release CCI 1.2.4 Features	**Closed**	**2013-12-11**	**2013-12-14**

History

#1 - 2013-12-17 05:17 - Rob Nahf

- Description updated

#2 - 2013-12-17 14:33 - Chris Jones

Rob, thanks for tracking this down. I'll chat with Ben and Jing about it, since this isn't as straight forward as I was thinking, because we need the local autogen id to be the same across the 3 CNs, even for DATA objects that aren't there.

BTW, edu.ucsb.nceas.metacat.dataone.hazelcast.HazelcastService implements the HZ EventListener interface, and updates the systemmetadata table on entry events.

#3 - 2013-12-17 16:07 - Rob Nahf

- Description updated

#4 - 2013-12-17 17:00 - Rob Nahf

It seems like the HZeventListener wasn't picking up the cn.registerSystemMetadata (entryAdded event) from the sync of the newly created DATA object.

In enrtyUpdated(), there's a conditional that only does a saveLocally() if the Member is not the owner (or vice versa?) But that seems like it wouldn't work for DATA format objects, because they are not on the local datastore to begin with. Maybe I'm interpreting that wrong...

#5 - 2013-12-19 15:53 - Chris Jones

- Status changed from New to In Progress

- Milestone changed from None to CCI-1.2

#6 - 2013-12-19 17:37 - Chris Jones

- Status changed from In Progress to Testing

Updated Metacat's CNodeService.delete() and .archive() methods. Will push to the 2.3 branch and will test this code.

See https://projects.ecoinformatics.org/ecoinfo/issues/6315

#7 - 2013-12-20 15:17 - Chris Jones

- Status changed from Testing to Closed

The issue with DATA object system metadata not being handled correctly on the CN is fixed. However, testing illuminated another bug with serialVersion, so I'll open a separate ticket.