Infrastructure - Story #3352

Production CNs are out of sync in object count

2012-10-22 17:03 - Chris Jones

Status:	Closed		**Start date:**	2012-10-23
Priority:	Normal		**Due date:**	2013-01-05
Assignee:	Ben Leinfelder		**% Done:**	100%
Category:	Environment.Production		**Estimated time:**	0.00 hour
Target version:	Sprint-2012.50-Block.6.4			
Story Points:				

Description

In the production environment, the CNs are slightly out of sync with regard to object count. cn-ucsb-1 is reporting 328691 objects, whereas both cn-orc-1 and cn-unm-1 are reporting 328448 objects. We need to figure out which pids are not in sync and why ucsb has fallen short.

Subtasks:

Task # 3357: Address ISet iterator bug that only iterates over a subset of the ISet	**New**

History

#1 - 2012-10-22 20:10 - Ben Leinfelder

An example (which probably applies broadly to the 200+ "missing" pids):
EML file guid 'nceas.988.9' was inserted into UCSB. For some reason there is no systemMetadata entry for it in the backing table, but it is mapped to docid 'autogen.2012101220031313139.1' in the identifier table.
ORC and UNM have systemMetadata for this guid (so HZ replication has worked) but they also have the docid as a separate object in the xml_documents table -- there from the Metacat replication of the EML file.

#2 - 2012-10-22 20:16 - Ben Leinfelder

Another pid with the same issue:
'kgordon.35.40' -> 'autogen.2012101713571339641.1'

#3 - 2012-10-22 21:48 - Chris Jones

At some point UCSB dropped from the Hazelcast cluster, but was still communicating via Metacat replication. Some objects were created on UCSB, but the call to saveLocally() looks to have failed, even though it succeeded on the other two CNs (via seeing the hz event). Ben created a short list of PIDs that were on UNM and ORC, but not UCSB, and so we'll restart UCSB to get the object counts to resync. The process will be:

1) Shut down d1-processing on cn-ucsb-1
2) Remove cn-ucsb-1 from the DNS round robin
3) Shut down d1-index-processing and d1-index-generator on cn-ucsb-1
4) Restart Tomcat on cn-ucsb-1
5) cn-ucsb-1 should resync it's missing identifiers in the system metadata tables by pulling them from the other two CNs
6) Restart d1-index-* on cn-ucsb-1
7) Restart d1-processing on cn-ucsb-1
8) Manually fix the incrrect guids on cn-orc-1 and cn-unm-1 for the short list of pids via SQL
9) Add cn-ucsb-1 back into the DNS round robin

Then,
10) Look through the cn-ucsb-1 logs to try to see why it dropped from the cluster in the first place.

#4 - 2012-10-22 22:34 - Ben Leinfelder

Numbers (1-4) and (8) are done. Now waiting for the resynch to complete on UCSB. I am skeptical it is in the cluster given the large number of statements like this flying by in the knb.log:
knb 20121022-22:34:11: [DEBUG]: Adding missing hzIdentifiers key: resourceMap_SRKX00_XXXIBTNXMBR11_20080505.50.2
[edu.ucsb.nceas.metacat.dataone.hazelcast.HazelcastService]

#5 - 2012-10-22 22:37 - Ben Leinfelder

- File correct_cn_20121022.sql added

- File correct_cn_20121022.sql added

#6 - 2012-10-23 17:21 - Chris Jones

- Tracker changed from Task to Story

- Due date set to 2012-10-27

- Status changed from New to In Progress

- translation missing: en.field_remaining_hours set to 0.0

#7 - 2012-10-23 17:29 - Chris Jones

We confirmed that cn-ucsb-1 was in the cluster, but there was a problem with it seeing all pids in the hzIdentifiers ISet. See subtask #3357.

#8 - 2012-11-08 00:38 - Chris Jones

We're now seeing that Hazelcast connections are frequently being dropped:

WARNING: /64.106.40.7:5701 [DataONE] hz.1.InThread Closing socket to endpoint Address[160.36.13.152:5701], Cause:java.io.EOFException

This looks like a known 1.9.X issue, and the recommended fix is to upgrade to Hazelcast 2.x

#9 - 2012-12-12 16:54 - Chris Jones

- Due date changed from 2012-10-27 to 2013-01-05

- Target version changed from Sprint-2012.41-Block.6.1 to Sprint-2012.50-Block.6.4

#10 - 2013-03-01 18:32 - Ben Leinfelder

- Status changed from In Progress to Closed

Counts are now synched.

Files

correct_cn_20121022.sql	6.03 KB	2012-10-22	Ben Leinfelder