

Infrastructure - Task #7466

Some objects not accessible on the CN via REST API

2015-11-04 18:41 - Bryce Mecum

Status:	In Progress	Start date:	2015-11-04
Priority:	Normal	Due date:	
Assignee:	Chris Jones	% Done:	30%
Category:	Metacat	Estimated time:	0.00 hour
Target version:		Story Points:	
Milestone:	None		
Product Version:	*		
Description			
While doing other work, I noticed that a good number (not sure how many) of objects listed on the CN's Solr index are not accessible via the REST API get() and resolve() methods. Instead of returning the object, they return a NotFound error.			
To reproduce,			
1. Visit https://cn.dataone.org/cn/v1/query/solr/?fl=identifier,title,authoritativeMN,datasource&q=formatType:METADATA+AND+-obsoleteBy:*&rows=100&start=0			
2. Pick a PID from the query result, e.g.			
<ul style="list-style-type: none">• knb-lter-cap.148.9• CLOEBDMETADATA.10242013.1			
1. Attempt to resolve() or get() the object via the REST API like: https://cn.dataone.org/cn/v1/object/CLOEBDMETADATA.10242013.1			
2. Receive a NotFound error instead of the object.			
Notes:			
In IRC, Skye noticed that the objects can be retrieved via their respective MN so it appears this issue may be a Metacat replication issue.			

History

#1 - 2015-11-04 18:46 - Skye Roseboom

Looks like a partial metacat replication between the CN:

does not work: <https://cn-ucsb-1.dataone.org/cn/v1/object/knb-lter-cap.148.9>
works: <https://cn-orc-1.dataone.org/cn/v1/object/knb-lter-cap.148.9>
works: <https://cn-unm-1.dataone.org/cn/v1/object/knb-lter-cap.148.9>

same with CLOEBDMETADATA.10242013.1
does not work: <https://cn-ucsb-1.dataone.org/cn/v1/object/CLOEBDMETADATA.10242013.1>
works: <https://cn-orc-1.dataone.org/cn/v1/object/CLOEBDMETADATA.10242013.1>
works: <https://cn-unm-1.dataone.org/cn/v1/object/CLOEBDMETADATA.10242013.1>

Can we trigger metacat replication audit between the CN so these objects and others are added to cn where missing?

#2 - 2015-11-04 18:47 - Skye Roseboom

- Category set to Metacat

#3 - 2015-11-04 18:48 - Skye Roseboom

- Assignee set to Chris Jones

Hey Chris - this looks like an issue that can be repaired by invoking the metacat replication repair service (Im not sure what the process is called). Can you trigger the metacat replication audit in production?

#4 - 2015-11-05 16:11 - Chris Jones

- Status changed from New to In Progress

- % Done changed from 0 to 30

I've forced Metacat replication manually by:

- Log into <https://cn-ucsb-1.dataone.org/metacat/admin>
 - Choose 'Replication Configuration' > 'Reconfigure Now'
 - Under 'Replicate Now', choose 'Get All'
- Repeat this for cn-orc-1 and cn-unm-1

I've also reset the timed replication from 172800000 (48 hrs) to 7200000 (2 hrs). We've discussed this in the past, and decided that the CNs should force replication on a much shorter time frame than the default in order to compensate for any network glitches that cause replication events to be missed. I'll be sure this is in the CN install/upgrade documentation.

Once replication has finished pulling from the other CNs, I'll check back to see if these NotFound errors are resolved.