Infrastructure - Bug #1514

Replication errors in ObjectFormat

2011-05-02 16:37 - Robert Waltz

Status:	Closed	**Start date:**	2011-05-02
Priority:	Urgent	**Due date:**	
Assignee:	Ben Leinfelder	**% Done:**	100%
Category:	Metacat	**Estimated time:**	0.00 hour
Target version:	Sprint-2011.18-Block.3		
Milestone:	CCI-0.5	**Story Points:**	
Product Version:	*		

Description
Replication does not pass objectFormat correctly. Upon initial create, objectFormat is set accurately. Upon replication to a subsequent coordinating not, the default application/octet-stream, is set as objectFormat

History

#1 - 2011-05-02 17:25 - Robert Waltz

- Target version set to Sprint-2011.18-Block.3

- Milestone set to CCI-0.5

#2 - 2011-05-02 20:33 - Robert Waltz

- File cn-unm-1_replicationErrorsFromLog.txt added

- Priority changed from Normal to Urgent

adding log messages from one replicated object

#3 - 2011-05-03 21:02 - Ben Leinfelder

Here's how I believe the servers are arranged with their synchronization/replication arrows drawn in:
knb-mn --> cn-ucsb-1

knb-test-1 --> cn-unm-1

cn-ucsb-1 <--> cn-unm-1

Tracking down this 'jwim1.5' data file that is referenced in the metadata document 'jwim1.6'....

On the originating member node knb-mn the content is actually the error message (http://knb-mn.ecoinformatics.org/knb/metacat/jwim1.5.1/default)!
<?xml version="1.0"?>

the requested docid 'jwim1.5' does not exist

This makes sense now that I know knb-mn was populated from the KNB where the jwim1.5 data file does not exist either. MetacatPopulator "reads" the KNB data from a URL, and that error message is considered the "content". Not the best, but It really shouldn't matter what is in this data file.

On the coordinating node cn-ucsb-1 this data file exists - but the *content* is blank (http://cn-ucsb-1.dataone.org/knb/metacat/jwim1.5.1/default). Coordinating nodes shouldn't have data (should only point to the member node that should have the data, correct?)

On the replicated CN (cn-unm-1) the data file exists and the content is still blank (http://cn-unm-1.dataone.org/knb/metacat/jwim1.5.1/default)

On the member node (knb-test-1) the document does not exist and gives this error message when attempting to read it (
[http://knb-test-1.dataone.org/knb/metacat/jwim1.5.1/default):](http://knb-test-1.dataone.org/knb/metacat/jwim1.5.1/default):
<?xml version="1.0"?>

the requested docid 'jwim1.5' does not exist

#4 - 2011-05-03 23:12 - Ben Leinfelder

- Status changed from New to Closed

- Assignee set to Ben Leinfelder

- % Done changed from 0 to 100

Metacat replication was not consistently using "object_format" when passing this metadata item from one instance to another. Now it is referred to as "object_format" across the board and the object format from the source Metacat is correctly set on the target Metacat. I should note that the source Metacat has my data files as "application/octet-stream" even in cases when it is a text/csv file. (Or am I expecting too much of object format at this point in the game?)

#5 - 2011-05-03 23:15 - Ben Leinfelder

I also want to point out that there were quite a few other [unrelated] errors being thrown during replication when I was working on this bug. Replication seemed to be processing the same documents at least twice which resulted in duplicate primary key constraint violations during insert. It doesn't appear that my target Metacat was hurt by this, but it does make for ugly logs and inefficient processing.

Files

cn-unm-1_replicationErrorsFromLog.txt	12.2 KB	2011-05-02	Robert Waltz