

Metacat - Bug #5536

Restore replicated data files that are 0 bytes

11/09/2011 12:23 PM - ben leinfelder

Status:	Resolved	Start date:	11/09/2011
Priority:	Normal	Due date:	
Assignee:	ben leinfelder	% Done:	0%
Category:	metacat	Estimated time:	0.00 hour
Target version:	2.0.0	Spent time:	0.00 hour
Bugzilla-Id:	5536		
Description			
Due to a replication bug, some data files were not replicated fully. These should be tracked down and restored.			
Related issues:			
Blocks Metacat - Bug #5519: Replicated data files are 0 bytes		Resolved	10/26/2011

History

#1 - 11/09/2011 12:47 PM - ben leinfelder

Matt suggests an alternative approach: delete the files and the DB records for those files so that the next timed replication will pick them up. This should certainly be tested before trying.

It may not work because I believe timed replication relies on the update date for the document when constructing a diff list and as far as I know there's not a "tell me every docid you have and I'll see if you're missing some that I have".

#2 - 12/15/2011 12:22 PM - ben leinfelder

Current proposal for upgrade script solution:

~~find all data files in the configured data directory for Metacat~~

~~delete those docids from the xml_documents and/or xml_revisions tables~~

~~delete those files from the filesystem~~

~~set xml_replication entry for that table (in remote Metacat) to have never run timed replication~~— this should force those data files to be re-replicated to the targets that had 0 byte data files.

Note that it requires extensive re-replication. Fine for small installations like PPBio or PELD, but not for PISCO or LTER.

Alternative, manual process:

-on replication source server: create tar file of all non-0 byte length files

-untar the archive on the target server, overwriting any existing (0 byte) files.

*repeat the process in the other direction if this was 2-way replication.

#3 - 01/03/2012 10:59 AM - ben leinfelder

I created another "upgrader" that will remove the empty replicated data files so that they can be re-replicated using timed replication.

#4 - 01/10/2012 09:23 AM - ben leinfelder

By removing replicated data files in my test environment (DEMO3 <-> fred) and removing the entries from the xml_documents and xml_revisions tables, timed replication successfully re-replicated the newly-missing data files.

I'm pretty happy with this solution as it does not require manual transfer of the zero-length replicated data files from source to target.

#5 - 03/27/2013 02:30 PM - Redmine Admin

Original Bugzilla ID was 5536