

Metacat - Bug #5519

Replicated data files are 0 bytes

10/26/2011 06:31 PM - ben leinfelder

Status:	Resolved	Start date:	10/26/2011
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:	5519		
Description			
<p>We set up 2-way replication between PPBio and PELD (Brazilian Metacat instances) and have since seen that the replicated data files have no content on the target server.</p> <p>-----</p> <p>Hi Ben,</p> <p>today I noticed something in some data in the replication between PELD and PPBio2.</p> <p>When I open a datafile that was originally inserted in Metacat through PELD, it is ok, but when I open the same datafile in PPBio2, the file is empty.</p> <p>The same thing happens the other way around.</p> <p>Ex: This package was inserted originally in PELD:</p> <p>http://peld.inpa.gov.br/knb/metacat?action=read&qformat=peld&sessionid=0&docid=menger.17.3</p> <p>But in PPBio2 the data file is empty:</p> <p>http://ppbio2.inpa.gov.br/knb/metacat?action=read&qformat=ppbio&sessionid=0&docid=menger.17.3</p> <p>File inserted originally on PPBio2:</p> <p>http://ppbio2.inpa.gov.br/knb/metacat?action=read&qformat=ppbio&sessionid=0&docid=menger.358.3</p> <p>Empty file: http://peld.inpa.gov.br/knb/metacat?action=read&qformat=peld&sessionid=0&docid=menger.358.3</p>			
Related issues:			
Blocked by Metacat - Bug #5536: Restore replicated data files that are 0 bytes			Resolved 11/09/2011

History

#1 - 10/31/2011 12:34 PM - ben leinfelder

I set up one-way replication from bespin.nceas.ucsb.edu --> fred.msi.ucsb.edu.

Using the Tomcat-only certificate generation/exchange approach, the data file was successfully replicated (with its content) to fred from bespin.

#2 - 10/31/2011 01:25 PM - ben leinfelder

Switching bespin.nceas.ucsb.edu to use Apache for the SSL handling still left the data replication working as expected (the data is in the replicated file).

#3 - 10/31/2011 02:06 PM - ben leinfelder

Now with BOTH servers using Apache certificate/SSL mechanisms the replication is still successful.

bespin.nceas.ucsb.edu is running a 1.10.0 variant with the semtools plugin and fred is running a pretty up-to-date version of the same (we recently changed trunk to use 2.0.0 versioning).

So this looks like a problem that exists in the 1.9.5 tag/branch and not currently in trunk -- will try to track down where it might have been resolved.

#4 - 10/31/2011 02:18 PM - ben leinfelder

There's no difference between 1.9.5 tag and trunk for the ReplicationService.handleGetDataFileRequest() method -- maybe there's a configuration difference that is causing this for peld->knb?

#5 - 10/31/2011 05:13 PM - ben leinfelder

After installing Metacat 1.9.5 on fred.msi.ucsb.edu I was able to see the failure point. It was something that I fixed in trunk a long time ago (probably during my character encoding enhancements) but never made it into our release branches. I'm not sure how old an issue this is for us, but I can't imagine it ever worked coded the way it is in 1.9.5 (which is an evolution of 1.9.3 when it first branched from trunk).

When ReplicationServlet handles any request it first gets a PrintWriter from the response object. But when we try to send data to the response by getting its OutputStream, we throw an IllegalStateException -- only one of those can be called on the same response object.

stack trace:

```
-----
knb 20111101-00:02:26: [DEBUG]: ReplicationServlet.handleGetOrPost - Action "readdata" accepted for server:
bespin.nceas.ucsb.edu/knb/servlet/replication [ReplicationLogging]
knb 20111101-00:02:26: [ERROR]: Servlet.service() for servlet replication threw exception
[org.apache.catalina.core.ContainerBase.[Catalina].[localhost].[/knb].[replication]]
java.lang.IllegalStateException: getWriter() has already been called for this response
at org.apache.catalina.connector.Response.getOutputStream(Response.java:570)
at org.apache.catalina.connector.ResponseFacade.getOutputStream(ResponseFacade.java:181)
at edu.ucsb.nceas.metacat.replication.ReplicationServlet.handleGetOrPost(ReplicationServlet.java:160)
at edu.ucsb.nceas.metacat.replication.ReplicationServlet.doGet(ReplicationServlet.java:77)
at javax.servlet.http.HttpServlet.service(HttpServlet.java:690)
at javax.servlet.http.HttpServlet.service(HttpServlet.java:803)
at org.apache.catalina.core.ApplicationFilterChain.internalDoFilter(ApplicationFilterChain.java:269)
at org.apache.catalina.core.ApplicationFilterChain.doFilter(ApplicationFilterChain.java:188)
at org.vfny.geoserver.filters.SetCharacterEncodingFilter.doFilter(SetCharacterEncodingFilter.java:122)
at org.apache.catalina.core.ApplicationFilterChain.internalDoFilter(ApplicationFilterChain.java:215)
at org.apache.catalina.core.ApplicationFilterChain.doFilter(ApplicationFilterChain.java:188)
at org.apache.catalina.core.StandardWrapperValve.invoke(StandardWrapperValve.java:210)
at org.apache.catalina.core.StandardContextValve.invoke(StandardContextValve.java:174)
at org.apache.catalina.authenticator.AuthenticatorBase.invoke(AuthenticatorBase.java:433)
at org.apache.catalina.core.StandardHostValve.invoke(StandardHostValve.java:127)
at org.apache.catalina.valves.ErrorReportValve.invoke(ErrorReportValve.java:117)
at org.apache.catalina.core.StandardEngineValve.invoke(StandardEngineValve.java:108)
at org.apache.catalina.connector.CoyoteAdapter.service(CoyoteAdapter.java:151)
at org.apache.jk.server.JkCoyoteHandler.invoke(JkCoyoteHandler.java:200)
at org.apache.jk.common.HandlerRequest.invoke(HandlerRequest.java:283)
at org.apache.jk.common.ChannelSocket.invoke(ChannelSocket.java:773)
at org.apache.jk.common.ChannelSocket.processConnection(ChannelSocket.java:703)
at org.apache.jk.common.ChannelSocket$SocketConnection.runIt(ChannelSocket.java:895)
at org.apache.tomcat.util.threads.ThreadPool$ControlRunnable.run(ThreadPool.java:685)
at java.lang.Thread.run(Thread.java:662)
```

#6 - 11/09/2011 12:22 PM - ben leinfelder

This is fixed. We just need to figure out the files that have not replicated successfully. Any 0-length files on KNB and LTER are suspect. Then we will need to track down where they came from in order to get the content.

#7 - 03/27/2013 02:30 PM - Redmine Admin

Original Bugzilla ID was 5519