

Metacat - Bug #7156

Solr index still keep the `obsoletedBy` field even though it was removed from the system metadata and reindexed

11/10/2016 03:14 PM - Jing Tao

Status:	Resolved	Start date:	11/10/2016
Priority:	Normal	Due date:	
Assignee:	Jing Tao	% Done:	0%
Category:	index	Estimated time:	0.00 hour
Target version:	2.8.1	Spent time:	0.00 hour
Bugzilla-Id:			
Description			
<p>Chris reported the issue: we're trying to reindex a pid in Metacat, and can't seem to get it to drop the <code>`obsoletedBy`</code> attr. I changed the system metadata in hazelcast to remove that field (and confirmed it in postgres). After logging in and issuing <code>`action=reindex&pid=urn:uuid:f22d78a2-719a-4704-856b-cc02fa803290`</code>, I see the reindex happen in catalina.out. But then when checking it, <code>`obsoletedBy`</code> is still there. Do we have a way to remove the entry from Solr and start from scratch?</p> <p>or is there a setting in the indexer that's just merging, and not deleting attrs?</p> <p>this is on arcticdata.io</p>			

History

#1 - 11/16/2016 10:09 PM - Jing Tao

- Target version changed from 2.9.0 to 2.8.1

#2 - 11/17/2016 11:09 AM - Jing Tao

I tried to use curl command to call the `updateSystemmetadata` to remove the "obsoletedBy" field. But I got the error:
<?xml version="1.0" encoding="UTF-8"?><error detailCode="4869" errorCode="400" name="InvalidRequest">
<description>The request is trying to reset the obsoletedBy field in the system metadata of the object jing-tao.15.1. This is illegal since the obsoletedBy filed is set, you can't change it again.</description>
</error>

Seems we can't remove it regularly. I have to write a hazelcast client class to do it (it is a hack!)

#3 - 12/09/2016 04:56 PM - Jing Tao

- Status changed from New to Resolved

It turns out that our solr index merge mechanism had an issue. It always copied the fields from the existing solr doc in the solr server to overwrite our new generated solr doc. In this case, the new generated doc doesn't have this `obsoletedBy` field, but the existing solr doc did have one. So it copied the the existing one to the new one.

Our solution is: for the system metadata fields, we don't merge it.

It works great.