

## Metacat - Bug #4405

### PISCO, KNB and LTER have different query result for SBC datasets.

09/21/2009 12:22 PM - Jing Tao

<b>Status:</b>	Resolved	<b>Start date:</b>	09/21/2009
<b>Priority:</b>	Normal	<b>Due date:</b>	
<b>Assignee:</b>	Chris Jones	<b>% Done:</b>	0%
<b>Category:</b>	metacat	<b>Estimated time:</b>	0.00 hour
<b>Target version:</b>	2.0.0	<b>Spent time:</b>	0.00 hour
<b>Bugzilla-Id:</b>	4405		

#### Description

On Mon, Sep 21, 2009 at 10:18 AM, Margaret O'Brien <[mob@msi.ucsb.edu](mailto:mob@msi.ucsb.edu)> wrote:

Hi metacat dev -

When I hit these 3 metacats with the query below, I get 3 different resultsets.

<http://knb.ecoinformatics.org/knb/metacat> (returns 13 datasets, some older rev numbers)

<http://metacat.lternet.edu/knb/metacat> (returns 14 datasets, with newer revision numbers)

<http://data.piscoweb.org/catalog/metacat> (returns 0 datasets)

All sbc datasets are uploaded to data.piscoweb.org, and are replicated to the other 2. I believe it is a one-way replication to knb, and from there two-way knb<->lter, but I could be wrong.

thought you might want to know...

margaret

```
<pathquery version="1.2">
<returndoctype>eml://ecoinformatics.org/eml-2.1.0</returndoctype>
<returndoctype>eml://ecoinformatics.org/eml-2.0.1</returndoctype>
<returndoctype>eml://ecoinformatics.org/eml-2.0.0</returndoctype>
<returnfield>eml/dataset/title</returnfield>
<returnfield>eml/dataset/creator/individualName/surName</returnfield>
<returnfield>eml/dataset/creator/organizationName</returnfield>
<returnfield>eml/dataset/dataTable/entityName</returnfield>

<returnfield>eml/dataset/dataTable/physical/distribution/online/url</returnfield>
<querygroup operator="INTERSECT">
<queryterm casesensitive="false" searchmode="contains">
<value>reed</value>
<pathexpr>creator/individualName/surName</pathexpr>
</queryterm>
<queryterm casesensitive="false" searchmode="starts-with">
<value>knb-lter-sbc</value>
<pathexpr>@packageId</pathexpr>
</queryterm>
</querygroup>
</pathquery>
```

Hi Margaret,

This is indeed a problem. The difference between nceas and lternet servers may be due to failures in indexing. Its not clear at all why the pisco server would return 0 though. Do you know what the right answer should be (e.g., how many results should be returned, and what are the docids (including revision that should match)?)

Matt

SBC datasets are uploaded to the PISCO metacat, and I rely on replication to get them to LTER. But I have not queried PISCO's metacat since it upgraded to the latest version. I have my own catalog application (I don't use the index.jsp), and only use metacat for the storage and dataset display. So I did not see this problem till I created a search-box to directly query metacat for datasets.

I've asked Chris about the 0 hits - but I think he was/is on vacation. I am attaching an email (to chris) that includes the log from a recent restart of that tomcat. It has a lot of warnings and errors in it.

thanks,  
Margaret

The response from the lter metacat is correct: there are 14 datasets with creator=Reed, with these ids:

knb-lter-sbc.13.11  
knb-lter-sbc.21.8  
knb-lter-sbc.24.7  
knb-lter-sbc.14.8  
knb-lter-sbc.15.7  
knb-lter-sbc.17.12  
knb-lter-sbc.18.8  
knb-lter-sbc.19.6  
knb-lter-sbc.25.2  
knb-lter-sbc.26.1  
knb-lter-sbc.27.1  
knb-lter-sbc.28.2  
knb-lter-sbc.29.1  
knb-lter-sbc.30.1

I checked the KNB metacat tables, and all 14 of these documents are present in the xml\_documents table, with the proper revisions. Which strengthens my earlier thought that it is only the index that is not updated.

Matt

Since we replicate eml in this way PISCO -> KNB -> LTER and LTER has correct packages, this means replication working fine. So I agree with matt (on irc), this may be caused by indexing issue we just introduced. However, both LTER and KNB should run the same version of metacat (1.9.1). I couldn't understand why they have different behaviors.

Thanks,

Jing

**Related issues:**

Blocked by Metacat - Bug #4551: performance enhancement through index reduction

**In Progress** 11/13/2009

**History**

**#1 - 09/21/2009 12:23 PM - Jing Tao**

This bug may relate to bug 4367

**#2 - 09/24/2009 01:36 PM - Jing Tao**

After rebuilding the index (through metacat servlet API), margaret confirmed that knb and lter have the same (correct) list.

Mike, could we schedule a time to discuss the issue?

**#3 - 11/09/2011 12:44 PM - ben leinfelder**

Running the query against the PISCO server returned the expected 14 records.

#4 - 03/27/2013 02:26 PM - Redmine Admin

Original Bugzilla ID was 4405