

## Metacat - Bug #2153

### Metacat Performance: Add/drop indices on key columns

07/13/2005 06:14 PM - Saurabh Garg

<b>Status:</b>	Resolved	<b>Start date:</b>	07/13/2005
<b>Priority:</b>	Immediate	<b>Due date:</b>	
<b>Assignee:</b>	Jing Tao	<b>% Done:</b>	0%
<b>Category:</b>	metacat	<b>Estimated time:</b>	0.00 hour
<b>Target version:</b>	1.8	<b>Spent time:</b>	0.00 hour
<b>Bugzilla-Id:</b>	2153		
<b>Description</b>			
From Matt's email...			
DB indices aren't created on some key columns such as xml_nodes.nodedata (originally because indices weren't supported on fields that big). This means a full table scan of the 6.5M records in xml_nodes for every query. We need to inspect for missing indices and create them, and need to consider the effect of UPPER on these indices in case-insensitive queries and be sure to create corresponding UPPER indices too. Once the indices are in place we need to check EXPLAIN to be sure the optimizer is using the indices instead of doing table scans (see <a href="http://www.postgresql.org/docs/faqs.FAQ.html#4.7">http://www.postgresql.org/docs/faqs.FAQ.html#4.7</a> ). Also, indices slow down inserts and updates, so any that aren't in use should be dropped, which will take some analysis of the running DB (but this is low priority comparatively speaking).			
<b>Related issues:</b>			
Blocked by Metacat - Bug #1879: Metacat Performance: Summary		<b>New</b>	<b>01/18/2005</b>

#### History

##### #1 - 07/20/2005 01:07 PM - Saurabh Garg

Indexes can also be added to docid and nodetype in xml\_nodes.....

##### #2 - 12/08/2005 10:10 AM - Saurabh Garg

Not enough time to do this for 1.6 release. The performance is fine for now. We can take this up again in 1.7 release.

##### #3 - 09/27/2006 05:54 AM - Saurabh Garg

For each of the queries, EXPLAIN needs to be done. I have noticed that for some of the queries, full table scan is being done - even though we have a index created for that particular column. I have noticed this in Postgres - hence this needs to be investigated.

##### #4 - 11/02/2007 09:54 AM - Jing Tao

Besides adding indices mentioned in bug 1879:

```
xml_path_index on upper(nodedata)
xml_access    on lower(principal_name)
xml_access    on perm_type
xml_access    on permission
xml_access    on perm_order
xml_access    on subtreeid
xml_documents on lower(user_owner)
```

We added more indices in order to improve performance:

```
CREATE INDEX xml_index_idx2 ON xml_index (docid);
CREATE INDEX xml_nodes_idx4 ON xml_nodes (docid);
CREATE INDEX xml_documents_idx3 ON xml_documents (rootnodeid);
CREATE INDEX xml_index_idx3 ON xml_index (nodeid);
```

We also changed the data type to TEXT for following columns:

```
nodedata in xml_nodes;
nodedata in xml_nodes_revision;
queryresult_string in xml_queryresult;
nodedata in xml_path_index;
```

path in xml\_index;

**#5 - 11/09/2007 04:03 PM - Jing Tao**

Move to 1.7.1

**#6 - 03/27/2013 02:19 PM - Redmine Admin**

Original Bugzilla ID was 2153