

Metacat - Bug #2152

Metacat Performance: Reduce size of xml_nodes table

07/13/2005 06:11 PM - Saurabh Garg

Status:	Resolved	Start date:	07/13/2005
Priority:	Normal	Due date:	
Assignee:	Saurabh Garg	% Done:	0%
Category:	metacat	Estimated time:	0.00 hour
Target version:	1.6	Spent time:	0.00 hour
Bugzilla-Id:	2152		

Description

xml_nodes is huge (6.8x106 records), and grows with every insert and update by many thousands of records. By design we don't want to search deleted docs nor old revisions anyway, so these could and should be moved to their own table (xml_nodes_revisions) that parallels the xml_revisions table. This would mean a massive reduction in the xml_nodes table size (xml_nodes has 6,885,702 records now, would reduce to 2,861,435 records) and a much slower growth rate for that table. This alone could speed up query time substantially.

History

#1 - 07/13/2005 06:35 PM - Saurabh Garg

While implementing this, have to check whats the limit on the sequence numbers that are being assigned to xml_nodes table. This will be a problem if the sequence number start rotating.

#2 - 07/20/2005 10:28 AM - Saurabh Garg

From Postgres documentation, the limits on the sequence...

"Sequences are based on bigint arithmetic, so the range cannot exceed the range of an eight-byte integer (-9223372036854775808 to 9223372036854775807). On some older platforms, there may be no compiler support for eight-byte integers, in which case sequences use regular integer arithmetic (range -2147483648 to +2147483647)."

From Oracle documentation, the limits on the sequence...

"an ascending sequence starts at one and increments up to the greatest allowed bigint value3, a negative sequence starts at -1 and decrements to the least allowed bigint value."

Currently on KNB we are at 7354460 for xml_nodes and 7354457 for xml_index. Hence long way to go.

#3 - 09/07/2005 01:24 PM - Saurabh Garg

This has been fixed. A new table is created in the Metacat schema where nodes from old document revisions and deleted documents are stored. This feature is currently being used by servers like ecoinfo2 and sbcddata and is running fine. Closing the bug.

#4 - 03/27/2013 02:19 PM - Redmine Admin

Original Bugzilla ID was 2152