

SAM Web services - Feature #9190

Modify schema to improve performance of parameter queries

06/12/2015 12:12 PM - Robert Illingworth

Status:	New	Start date:	06/12/2015
Priority:	Normal	Due date:	
Assignee:		% Done:	0%
Category:		Estimated time:	0.00 hour
Target version:	x Future Development	Spent time:	0.00 hour
Description			
<p>The current implementation of parameters is flexible, but the multiple self joins required to perform dimension queries can cause performance issues. Possible solutions:</p> <ol style="list-style-type: none">1) Use a table with many columns to store all string params for a file in a single row. This means that only a single pass is required to evaluate all the parameter constraints. Postgres stores sparse tables in a compact form, and each column can have a partial index with a non-null requirement; this makes the whole thing more compact than the current arrangement, which means it'll fit better in the memory cache. There may be issues with retrieving the metadata for a given file as it now needs to skip all the null values. The width of the table will also limit the potential maximum number of parameters.2) Use hstore or json data types to store the data directly. May take significantly more space, and querying on anything other than value equality may be slow. Which since we want case insensitive queries is a problem.			