

Metacat - Feature #5989

Track data download, view and citation statistics

05/24/2013 06:14 PM - Matt Jones

Status:	In Progress	Start date:	
Priority:	Normal	Due date:	
Assignee:	Peter Slaughter	% Done:	0%
Category:		Estimated time:	0.00 hour
Target version:	2.x.y	Spent time:	0.00 hour
Bugzilla-Id:			
Description			
<p>Currently the only usage stats we have in Metacat are the raw logs. This new service would provide several statistical reports in machine-readable format intended for efficient use on clients for building user interface displays that show those statistics.</p> <p>The service should include the following response statistics, and be extensible to add other tracked statistics as needed:</p> <ol style="list-style-type: none">1. Number of views (defined as number of times the metadata has been viewed on the web)2. Number of package downloads (needs definition)3. Size in bytes of package downloads4. Number of citations (implement in a second phase) <p>For each of these statistics, calling apps should be able to constrain the results to only include records matching:</p> <ol style="list-style-type: none">1. a PID or list of PIDs2. creator or list of creators (DN, or ORCID, or some amalgam -- to be discussed)3. a time range of access event (upload, download, etc.)4. ? spatial location of access event (upload, download, etc.)5. ? IP Address6. accessor or list of accessors (DN, or ORCID, or some amalgam, needs ACL -- to be discussed) <p>For each of these statistics, calling apps should be able to request the statistic aggregated by several specific facets, including the following (in order of importance):</p> <ol style="list-style-type: none">1. User (DN, or ORCID, or some amalgam -- to be discussed)2. Time range, aggregated to requested unit (day, week, month, year)3. ? Spatial range, aggregated to requested unit (to be discussed) <p>Intersections of these aggregated facets should also be possible, but are a lower priority than the facets alone. For example, when finished, one should be able to request the following reports, among others:</p> <ol style="list-style-type: none">1. {Views,Downloads,Bytes,Citations} for a given pid or list of pids2. {Views,Downloads,Bytes,Citations} by user (aggregates across pids)3. {Views,Downloads,Bytes,Citations} by month (aggregates across pids)4. {Views,Downloads,Bytes,Citations} by spatial location (aggregates across pids)5. {Views,Downloads,Bytes,Citations} for a given pid by month for a specific time range6. {Views,Downloads,Bytes,Citations} by user by month7. etc. <p>The download format (JSON?, XML?) should allow for an extended set of response variables, and an extendable set of aggregating facets. Need to discuss, but probably XML as that is DataONE's initial choice for all other services. Contemplate both if useful.</p> <p>The REST API for this service should be developed in the DataONE space, with intention of it being implementable by both other MNs and CNs in DataONE.</p>			
Subtasks:			
Task # 5990: Track downloads			New
Task # 5991: Track views			New
Task # 5992: Track citations			New
Task # 5993: Summarize and index statistics for fast access			New
Task # 5994: Create REST API for accessing statistics			New

Related issues:

Related to Metacat - Feature #6346: Make # READ events available in SOLR index

Resolved**01/03/2014**

Has duplicate MetacatUI - Feature #6289: Create usage statistics service

Rejected**History****#1 - 08/23/2013 03:56 PM - ben leinfelder***- Target version changed from 1.1.0 to 1.2.0***#2 - 10/18/2013 04:10 PM - ben leinfelder***- Target version changed from 1.2.0 to 1.3.0***#3 - 11/14/2013 08:26 AM - ben leinfelder***- Target version deleted (1.3.0)***#4 - 12/09/2013 11:12 AM - Lauren Walker***- Target version set to 1.5.0***#5 - 12/09/2013 02:36 PM - Matt Jones***- Description updated***#6 - 12/10/2013 11:25 AM - Matt Jones***- Target version deleted (1.5.0)**- Project changed from MetacatUI to Metacat***#7 - 12/10/2013 11:36 AM - Matt Jones***- Target version set to 2.5.0**- Assignee set to Peter Slaughter***#8 - 12/10/2013 12:10 PM - Matt Jones***- Description updated*

Updated description to clarify a few of the reports and filters.

#9 - 12/10/2013 01:32 PM - Matt Jones*- Description updated***#10 - 12/10/2013 01:35 PM - Matt Jones***- Description updated***#11 - 12/10/2013 03:02 PM - Matt Jones***- Target version changed from 2.5.0 to 2.4.0***#12 - 01/13/2014 09:01 AM - Peter Slaughter***- Status changed from New to In Progress***#13 - 01/30/2014 12:17 PM - ben leinfelder***- Target version changed from 2.4.0 to 2.5.0***#14 - 05/20/2014 10:56 AM - ben leinfelder**

I believe this is now all targeted for the CNs. We may consider moving tasks into the DataONE tracker as appropriate...

#15 - 11/03/2015 11:31 AM - ben leinfelder*- Target version changed from 2.5.0 to 2.5.1***#16 - 12/14/2015 02:03 PM - ben leinfelder***- Target version changed from 2.5.1 to 2.x.y*