

Kepler - Bug #5297

Workflow Run Manager - no indication of progress during preview runs download

02/07/2011 12:42 PM - Derik Barseghian

Status:	Resolved	Start date:	02/07/2011
Priority:	Normal	Due date:	
Assignee:	Derik Barseghian	% Done:	0%
Category:	workflow run manager	Estimated time:	0.00 hour
Target version:	reporting-2.3.0	Spent time:	0.00 hour
Bugzilla-Id:	5297		

Description

There are currently about 131 run-kars on kepler-dev. Switching to kepler-dev from the workflow run manager for the first time takes (including authentication), upwards of 30s for me, and I'm on a fast university connection. A busy icon shows during the wait, but it would be better if each preview run showed up in the WRM as it was downloaded, instead of waiting until all are downloaded. Also look into speeding up this process, e.g. check with Jing about caching the search results on metacat, look for bottlenecks (disk i/o with the tracker xml file?), etc.

History

#1 - 02/07/2011 12:57 PM - Matt Jones

Derik,

When we discussed this before, I think we concluded that the search time was largely due to Kepler issuing an EcoGrid get() request for every result record, which obviously doesn't scale. The Metacat search itself should be subsecond for up to thousands of records. Metacat needs to be configured properly for that to work though -- in that the karxml fields need to be indexed for search to be fast. Other than that, its probably a client-side issue in its decision to download all of the KARs.

#2 - 02/07/2011 02:36 PM - Derik Barseghian

This bug is just for fetching data needed to build "preview" WorkflowRuns (bits from the karXML, not the kars or entire karXMLs). I get all records in one query to metacat. I did some quick timing tests of WRM with kepler-dev (there are a few reasons why these shouldn't be considered exact, but they give a good rough idea) :

The ~30s that it currently takes to show all ~131 preview runs for a first time switch of WRM to kepler-dev breaks down roughly:

- 1) "anon" authentication via kepler gui: ~6s
- 2) query to fetch all data needed to build workflowRuns: ~4s
- 3) locally building runs, writing to provenance and tracker file: ~17s

For a subsequent Refresh action:

- 1) ~0s (already authenticated)
- 2) ~3s
- 3) ~0s

So I'll look into improving [#3](#), the big cost must be the write actions.

#3 - 02/08/2011 04:20 PM - Derik Barseghian

"fixed" at r27053.

Step [#3](#) now takes about 2s instead of 17.

Refactoring to show the preview runs as they're imported would take a fair bit of work, especially to avoid putting us right back at 17s+, and at 2s wait it's not worth it.

#4 - 03/27/2013 02:29 PM - Redmine Admin

Original Bugzilla ID was 5297