**caGridLogger**

**Version 0.5**

**26-Feb-2009**

**Ian Fore**

**NCI Center for Biomedical Informatics and Information Technology**

Document history

|  |  |
| --- | --- |
| **Date** | **Notes** |
| 26-Feb-2009 | Initial deployment to GForge |

## Current capabilities

Query all services by service name i.e. query all caTissueSuite, all caArraySvc, caNanoLabSvc, caTissueCore etc.

Query on a specific named grid – currently NCI production and training grids.

Specify a list of classes to be queried for a particular service type.

Log count of each class in a service instance.

Log time taken for query

Log any error/exception that occurred for a particular query.

Support easy display, reporting and analysis of results over time.

Code origin

The code is based on a couple of caGrid Example projects

See

<https://cabig-kc.nci.nih.gov/CaGrid/KC/index.php/Authentication_and_Authorization_for_Client_Developers>

and the section below on libraries

## Running the code

There are three main targets in the ant build file build.xml

**run**

This is the default target. It runs the class gridexamples.ListServices

This queries caGrid for services which match one or more service names specified in the build file. It then queries each service for a defined list of object types and returns the count of each object type. It uses a CQL count query to do so. The count query may fail or timeout for certain object types. This is recorded and logged to a database.

**runCP**

Runs a grid query to list Collection Protocols against all caTissueSuite grid services. This example is awaiting completion.

**runLogin**

This runs the example code provided by the caGrid Knowledge Center to demonstrate grid authentication and authorization.

## Database

A database is used by the application to

Log results

Maintain information about which classes should be queried for any particular model.

Maintain model metadata not (or not easily) available on caGrid.

The DBMS currently used is FileMaker Pro. A main reason for using FileMaker Pro is that it is convenient to create quick reports and a UI through which to control the aggregator.

The database used to date is included in CVS in the data subdirectory.

FileMaker Pro has limited JDBC functionality. However, that has not been a problem to date with the simple database structure required.

## Libraries

caGrid offers a dependency management system based on Ivy which is described in the caGrid documentation and here

<http://www.cagrid.org/wiki/CaGrid:How-To:DependOnCaGridLibraries>

That has not been used completely successfully on this project. Currently the set of jars are in the lib directory. Using Ivy will overwrite the directory and is therefore not yet recommended.

## To dos

Query for specific class attributes e.g. Collection protocol names – QueryCPs class is a work in progress to do this

Query for participant counts by Collection protocol – intention is that QueryCPs class will do this.

For a given service type allow more than one set of classes to be saved for query

Allow query by model type – tried this but it failed for some reason.