caNanoLab 1.2

Installation Guide

September 11, 2007

Table of Contents

Introduction to caNanoLab

Welcome to the cancer Nanotechnology Laboratory Analysis Bench (caNanoLab) 1.2 Installation Guide. caNanoLab is a Laboratory Information Management System (LIMS) designed to facilitate data sharing in laboratories performing nanoparticle assays. caNanoLab provides support for accessioning of samples (nanoparticles), execution of nanoparticle assays, and recording of assay results. Additionally, caNanoLab allows for annotation of nanoparticles with characterizations resulting from physical and *in vitro* nanoparticle assays, as well as sharing of these characterizations in a secure fashion.

As of release 1.1, the caNanoLab domain model has been caGrid enabled. In other words, a caNanoLab grid data service can be deployed and registered with caGrid 1.0 index server, allowing sharing of public characterizations across the caGrid. For more information, see https://cabig.nci.nih.gov/workspaces/Architecture/caGrid. The caNanoLab web application now allows remote searches against remote grid services hosting the caNanoLab data model.

Targeted Developer

Although the installation instructions have been simplified as much as possible, this installation guide is best suited for an experienced java developer who has some familiarity with the following J2EE and related technologies:

Important Background Knowledge

- Unix/Linux environment
- (Configuring environment variables; Installing Ant, JDK, Apache Tomcat and JBoss servers)
- Ant build scripts
- J2EE web application development using the Struts framework, Servlet/JSP's, JavaScript
- J2EE middle-ware technologies such as n-tier service-oriented architecture and software design patterns
- caGrid 1.0 infrastructure (for setting a caNanoLab grid data service)

General System Requirements

The following technologies power a caNanoLab web application:

Technologies

- Java Software Development Kit (JDK) version 5.0 http://java.sun.com/javase/downloads/index_jdk5.jsp
- JBoss version 4.0.5 http://labs.jboss.com/jbossas/downloads
- Jakarta Ant version 1.6.5 http://archive.apache.org/dist/ant/binaries/
- Oracle 9i Release http://www.oracle.com/technology/software/products/oracle9i/index.html

---

The caNanoLab web application has been tested within NCICB against JBoss servers hosted on Windows XP and RedHat Linux systems, and against Oracle 9i databases hosted on Sun Solaris systems, and Oracle 10g XE database hosted on Windows XP.

Download each of the tools listed in the bulleted list and follow the installation instructions provided with each respective product for your environment. Assistance from an Oracle 9i database administrator is expected to properly configure the Oracle database.

Grid Service

Setting up a caNanoLab grid service is optional, and it is recommended that you install the grid service only after you successfully install the web application and become familiar with submitting and searching data through the web application.

Grid Service Downloads

Once you have data that you would like share with the caBIG community and you intend to set up a caNanoLab grid service, download the following open source technologies that power a caNanoLab grid data service:

- The Globus Toolkit version 4.0.3 http://www.globus.org/
- Apache Tomcat version 5.0.28 http://tomcat.apache.org/

Obtaining the caNanoLab Source

Source Code

The caNanoLab web application source code is distributed as a zip file named `caNanoLab_1.2.zip` from the NCICB download center at http://ncicb.nci.nih.gov/download/.

There are no source updates related to the caNanoLab grid service. The caNanoLab grid service source code is distributed as a zip file named `caNanoLab_grid_1.1.zip` from the NCICB download center at http://ncicb.nci.nih.gov/download/.

Database Technology

Assumptions and Requirements

The caNanoLab source distribution `caNanoLab_1.2.zip` has been downloaded from the NCICB download site. An Oracle database has been set up on a system (local or remote) with an administrator account; you are familiar with the process of creating tablespaces, creating users and assigning default tablespaces and privileges to users, and importing data from an Oracle dump file.

For a Previous Installation

If you have installed caNanoLab release 1.1 or release 1.1.1 against an Oracle database and have associated production data in these schemas and you would like to continue to use the same schemas for caNanoLab release 1.2, review the following database initialization steps, then go directly to *Data Migration* on page *4*.

For a New Installation

If you are installing the caNanoLab application for the first time or want to install new schema for release 1.2, follow the steps below to set up the required schema objects and the seed data for release 1.2.

Follow these steps to initialize your system:

Step	Action

1	Log into the Oracle database as the database administrator and create a named tablespace with at least 1 GB in size. Create a user with this tablespace as the default tablespace and grant this user privileges to connect, create table, create view, create sequence and create trigger. Refer to Oracle documentation for details.
	
	Example: On a Windows XP system hosting an Oracle 10g XE database, you would issue the following commands at the SQL prompt:
	
	`SQL>create tablespace cananolabd datafile 'c:\oraclexe\oradata\xe\cananolabd.dbf' size 1024M permanent;`
	
	`SQL>create user cananoLab12 identified by go1234 default tablespace cananolabd;`
	
	`SQL>grant connect, create table, create view, create trigger, create sequence, unlimited tablespace to cananolab12;`

2	Extract the `caNanoLab_1.2.zip` to a location on your local system, for example, `C:\caNanoLab_1.2`. This location is referred as `<CANANOLAB_SOURCE>` throughout the document. Verify that a database dump file `caNanoLab_1.2.dmp` exists in the directory `<CANANOLAB_SOURCE>/db`.

Step	Action

3	Import the database dump file, `caNanoLab_1.2.dmp`, into the user schema created in step 1.
	
	Example: On a Windows XP system hosting an Oracle 10g XE database, you would issue the following commands at the DOS prompt:
	
	`C:\>cd c:\caNanoLab_1.2\db\1.2`
	
	`C:\>c:\oraclexe\app\oracle\product\10.2.0\server\BIN\imp cananolab12/go1234 file=caNanoLab_1.2.dmp log=caNanoLab_1.2.log full=y`

Verification

Once the database has been created, either through brand new data import or through data migration (described in the next section), verify that the following numbers of database objects are created:

- **Tables 101**
- **Views 1**
- **Triggers 25**
- **Sequences 12**

Example: Issue the following queries at the SQL prompt:

```
SQL>select count(*) from all_tables where owner='CANANOLAB12';

SQL>select count(*) from all_views where owner='CANANOLAB12';

SQL>select count(*) from all_triggers where owner='CANANOLAB12';

SQL>select count(*) from all_sequences where sequence_owner='CANANOLAB12';
```

Data Migration

NOTE:	
	If you are installing caNanoLab 1.2 for the first time or installing a new caNanoLab release 1.2 schema, you can skip this section.

If you have already installed caNanoLab release 1.1 or release 1.1.1 previously and have production data associated with the schemas and wish to keep the same schemas, you must perform a data migration to get the latest schema and seed data. During execution of the migration script, you will be prompted to enter a value for a variable named `appowner`, representing the owner of the caNanoLab 1.2 installation instance. Enter a value that is appropriate for your institute, for example, NCL, NCICB, WashU, MIT, Stanford, etc. This value is used as a prefix to the three default user groups (for example, NCL_Administrator, NCL_PI, NCL_Researcher) generated during the caNanoLab application start up. These users groups are essentials to the caNanoLab application role-based authorization. Details are described on page 10.

NOTE: 	If you have started previous releases of caNanoLab, you should already have these three user groups defined in the GROUP_NAME column of the table CSM_GROUP. The migration script will take the new `appowner` value you enter when running the migration script and make the appropriate updates in the database if it is necessary to use this new value.

Follow these steps to complete the required data migration:

Step	Action
1	Locate the data migration script `migration_script.sql` at the directory `<CANANOLAB_SOURCE>/db/1.2/data_migration/from_1.x`, where x=1 or 1.1, e.g. if you are migrating from release 1.1, use the script under the subdirectory `from_1.1`.

2	Go to the directory that contains the migration script and run the SQL script at the command line as the user for the caNanoLab schema. For example, on a Windows XP system hosting an Oracle 10g XE database, you would issue the following commands at the DOS prompt:
	
	`C:\>cd c:\caNanoLab_1.2\db\data_migration\from_1.1`
	
	`C:\>c:\oraclexe\app\oracle\product\10.2.0\server\BIN\sqlplus cananolab12/go1234 @migration_script.sql`

After data migration, refer to the Verification section on page 4 to verify that the migration has been successful. The same number of schema objects should be generated as described in that section.

caNanoLab Web Application

Assumptions and Requirements

The database has been installed and verified, as described on pages 3 and 4. Ant has been installed. The JBoss application server has been installed on a system (local or remote) and can be started at a designated port. JBoss install directory is referred as `<JBOSS_HOME>` in the document. The JBoss application server host URL is referred as `<SERVER_URL>`. It is assumed that the default configuration is used for deploying caNanoLab related web archive files, and it is recommended that you change the designated port from the default port 8080 to another unused port.

For example, in the case of JBoss 4.0.5, the default configuration is located at the directory `<JBOSS_HOME>/server/default`. The caNanoLab web archive file shall be deployed at the directory `<JBOSS_HOME>server/default/deploy`. For changing the default port 8080 and associated ports, locate the file `server.xml` at the directory `<JBOSS_HOME>/server/default/deploy/jbossweb-tomcat55.sar`. Replace all occurrences of 8080, 8009 and 8443 with 18080, 18009 and 18443, respectively. For details, please refer to the appropriate JBoss documentation.

Installation and Deployment

Follow these steps to install and deploy caNanoLab.

Step	Action
1	Edit the Ant build properties file `build.properties` at `<CANANOLAB_SOURCE>` by specifying values for the following five properties: a. `hostname`: The caNanoLab application hostname, for example, `localhost:18080/caNanoLab` b. `file.repository.dir`: A directory on the system that hosts the JBoss application server for storing uploaded files, for example, `c:\caNanoLab_Data`. **Note:** This directory should be writable by the user that starts the JBoss server, and this directory should be created prior to starting the application. c. `sample.prefix`: The prefix used for auto-accessioning samples, for example, NCL is used for sample accessions `NCL-1, NCL-2`, etc at the Nanoparticle Characterization Laboratory. d. `application.owner`: the owner of the caNanoLab release 1.2 installation instance, for example, NCL. **Note:** This value should be the same as the `appowner` value used during data migration, described on page 4. e. `grid.indexserver`: the grid index server URL hosting the caNanoLab grid data services, for example, `cagrid-index-qa.nci.nih.gov:8080`, the NCICB's QA caGrid index server.

Step	Action

2	Execute the Ant build script `build.xml` located at `<CANANOLAB_SOURCE>` with the default target `build-application` and the default `build.properties` file configured in step 1.
	
	Example: Issue the following commands to execute the Ant script:
	
	`C:\>cd c:\caNanoLab_1.2`
	`C:\>ant`
	
	Successful execution of the Ant script will create three deployable web archive war files in the directory `<CANANOLAB_SOURCE>/output/war`: `caNanoLab.war`, `upt.war` and `caNanoLabSDK.war`. Place these war files into the default JBoss application server deploy directory, `<JBOSS_HOME>/server/default/deploy`.

3	Copy the file `ojdbc14.jar` from `<CANANOLAB_SOURCE>/lib` folder and place it in the directory `<JBOSS_HOME>/server/default/lib/.`

CSM

 **Before You Start caNanoLab**	Before starting the JBoss application server and running the caNanoLab application, you must configure the server with required security configurations for user authentication and authorization. The caNanoLab web application relies on a modified version of NCICB CSM (Common Security Module) version 3.1 for user authentication and role-based authorization. For more information, see the CSM documentation at http://gforge.nci.nih.gov/frs/?group_id=12.

Follow these steps to complete the security configurations:

Step	Action

1	Copy the following files from the directory `<CANANOLAB_SOURCE>/conf/csm` and place them in the directory `<JBOSS_HOME>/server/default/conf`:
	
	`caNanoLab.csm.hibernate.cfg.xml`
	`ApplicationSecurityConfig.xml`

2	Edit the file `ApplicationSecurityConfig.xml` at `<JBOSS_HOME>/server/default/conf as follows`: • Replace the token `@ABSOLUTE_PATH@` with the value `<JBOSS_HOME>/server/default/conf`, where `<JBOSS_HOME>` should be replaced by the actual full path JBoss install directory name, `C:/jboss-4.0.5.GA`. **Note:** Either double back slashes \\ or a single forward slash / should be used as the file separator if working on Windows platform.

Step	Action
3	Copy the file `oracle-ds.xml` from the directory `<CANANOLAB_SOURCE>/conf` and place it in the directory `<JBOSS_HOME>/server/default/deploy`, if such a file does not already exists.
4	Edit the copied file `oracle-ds.xml` at `<JBOSS_HOME>/server/default/deploy` as follows: a. Replace the token `@DBSERVER@` with the URL or the IP address of the system that hosts the Oracle database, for example, ***127.0.0.1***. b. Replace the token `@DBNAME@` with the name of the database service to which to connect, for example, the value is ***xe*** for the default Oracle Database 10 gXE service. Check with your Oracle DBA for your database service name configurations. c. Replace the token `@DBUSER@` with the user schema name configured in step1 of the Database Technology instructions on page 3, for example, ***cananolab12***. d. Replace the token `@DBPASSWORD@` with the password associated with the value for the token `@DBUSER@`, for example, ***go1234***.
5	Locate the last `</application-policy>` tag in the file `login-config.xml` at `<JBOSS_HOME>/server/default/conf`, and insert the content of the file `login-config-segment.xml` at `<CANANOLAB_SOURCE>/conf/csm` into the file `<JBOSS_HOME>/server/default/conf/login-config.xml` just below the last `</application-policy>` tag.
6	Edit the inserted file `login-config.xml` at `<JBOSS_HOME>/server/default/conf` by doing similar replacements as in step 4: a. Replace the token `@DBSERVER@` with the URL or the IP address of the system that hosts the Oracle database, for example, ***127.0.0.1***. b. Replace the token `@DBNAME@` with the name of the database service to which to connect, for example, the value is ***xe*** for the default Oracle Database 10 gXE service. Check with your Oracle DBA for your database service name configurations. c. Replace the token `@DBUSER@` with the user schema name configured in step1 of section 1, for example, ***cananolab12***. d. Replace the token `@DBPASSWORD@` with the password associated with the value for the token `@DBUSER@`, for example, ***go1234***.
7	Locate the last </mbean> tag in the file `properties-service.xml` at `<JBOSS_HOME>/server/default/deploy`, and insert the content of the file `properties-service-segment.xml` at `<CANANOLAB_SOURCE>/conf/csm` into the file `<JBOSS_HOME>/server/default/deploy/properties-service.xml` just above the last </mean> tag.
8	Edit the inserted file `properties-service.xml` at `<JBOSS_HOME>/server/default/deploy` as follows: • Replace the token `@JBOSS_HOME@` with the value of the full path JBoss install directory name, for example, `C:/jboss-4.0.5.GA`. **Note:** Either double back slashes \\ or a single forward slash / should be used as the file separator if working on Windows platform.

Verification

Once the JBoss application server is correctly configured with the CSM settings, you can now start the JBoss application server and start the caNanoLab application. Open the URL http://<SERVER_URL>/caNanoLab/ (for example, http://localhost:18080/caNanoLab). You should see a disclaimer page. Click on **CLICKING HERE** to go to the welcome/login page.

UPT

Before users can log in to the caNanoLab application to submit and search data, you must first create their user accounts and assign them to the pre-defined user groups with the pre-defined roles to access the pre-defined protection groups. The caNanoLab application makes use of the NCICB's User Provisioning Tool (UPT), a separate web application, for user account management. The concepts of users, groups, roles, protection groups are defined according the CSM/UPT principles. See the CSM documentation at http://gforge.nci.nih.gov/frs/?group_id=12 for details on these concepts and the use of the UPT tool.

NOTE: 	As a part of the database seed data, a default user group Public has been created and has been assigned access to the SEARCH and the REMOTE SEARCH functions of the caNanoLab application. During the caNanoLab application start up, three more default users groups are auto-generated:
	
	`<application.owner>_Administrator,`
	`<application.owner>_Researcher`
	`<application.owner>_PI`
	
	Where the value for `<application.owner>` is specified in the file `build.properties` prior to building the application as described on page 6.

In release 1.2, during the caNanoLab application start up, automatic roles are assigned for these three default user groups as follows:

1. `<application.owner>_Administrator` is assigned access to the ADMINISTRATION function,
2. `<application.owner>_Researcher` is assigned access to the WORKFLOW and the INVENTORY functions, and
3. `<application.owner>_PI` is assigned access to the WORKFLOW, the INVENTORY and the SUBMIT functions.

UPT Example

The following steps illustrate an example use of the UPT tool to create a new user, to assign the user to be a caNanoLab administrator, and to assign the user to two pre-defined user groups.

Step	Action
1	Launch the UPT tool at http://<SERVER_URL>/upt and login as **superadmin/password**. Use **caNanoLab-upt** as the application name when prompted at the UPT log in. **Note:** The user **superadmin** was created as a part of the database setup.
2	Once logged into the UPT tool, follow these steps: a. Select **User > Create a New User**. b. Create a new user account named **admin**. c. Select **Application > Select an Existing Application**; click **Search**. d. Select **caNanoLab** from the application list. e. Click **View Details**, then **Associated Admins**. f. Assign this user to be an administrator for the caNanoLab application. g. Click **Update Association** to commit the change.
3	Log out of the UPT tool and log back in as **admin/admin**. Use **caNanoLab** as the application name when prompted at the UPT log in. **Note:** When a new user account is created, its initial default password is the same as its login name. Given that the UPT tool is lacking a function for users to manage their own passwords, users can only update their passwords within the caNanoLab application. See the Notes below on how to update user passwords.
4	Select **User > Select an Existing User**, and click **Search**. Select **admin** from the User list. a. Click **View Details**, then **Associated Groups**. b. Select **<application.owner>_Adminstrator** and **Public** groups from the pre-defined group list and assign them to the user. c. Click **Update Associations** to commit the change.

You can follow similar steps to create other application user accounts and to assign them to different users groups, according to your needs.

NOTES: 	• All users should be assigned to the Public group so everyone has access to the SEARCH and the REMOTE SEARCH functions. • A user can be assigned to multiple user groups, for example, **admin** is assigned to both <application.owner>_Administrator and Public groups. • The initial passwords for user accounts are set to be the same as their user accounts. When a user first logs into the caNanoLab application at http://<SERVER_URL>/caNanoLab, he/she would be prompted to change the initial password. Users can update their passwords at any time within the caNanoLab application.

For more information about how to use the UPT tool for managing user accounts, contact NCICB Application Support at ncicb@pop.nci.nih.gov and request that the caNanoLab technical team give you a demonstration of the UPT tool in the context of the caNanoLab application.

caNanoLab Grid Data Service

Assumptions and Requirements

Apache Tomcat has been downloaded and installed with the `CATALINA_HOME` environment variable set pointing to the Tomcat install directory. The Tomcat server URL is referred to as `<TOMCAT_SERVER_URL>` throughout the document.

Globus Toolkit has been downloaded and installed with the `GLOBUS_LOCATION` environment variable pointing to the Globus install directory. Apache Tomcat is assumed to be the grid service container.

The caNanoLab grid data service requires the caNanoLab web application war file `caNanoLab.war` be deployed, as well as caNanoLab SDK application service war file `caNanoLabSDK.war` be deployed to the same application server. Refer back to the caNanoLab web application installation and deployment section on page 6 for details on how to deploy these war files.

Installation and Deployment

Before deploying grid services, Tomcat must be pre-configured with a Globus web application that can host a grid service. Issue the following commands to complete the pre-configuration required for Tomcat to host a caNanoLab grid service:

```
C:>cd "%GLOBUS_LOCATION%"
C:>ant -f share\globus_wsrf_common\tomcat\tomcat.xml
deployTomcat -Dtomcat.dir="%CATALINA_HOME%" –
Dwebapp.name=wsrf-canano
```

This will generate a directory `<CATALINA_HOME>/webapps/wsrf-canano` with the required Globus webapp information. Note that by specifying a `webapp.name` property (in the above command) to have a different value than the default value `wsrf`, multiple caCORE SDK backed grid services can be deployed to the same Tomcat server and function properly and independently without interference. This circumvents the current known issue with multiple caCore SDK-backed grid services interfering with each other if deployed to the same Globus webapp `wsrf` under the same Tomcat server.

Once the Tomcat server is pre-configured, complete the following steps to create and deploy a caNanoLab grid data service.

Step	Action
1	Extract `caNanoLab_grid_1.1.zip` to a location on your local machine such as C drive. This location is referred as `<CANANOLAB_GRID_SOURCE>` throughout the document.
2	By default, the caNanoLab grid data service is registered to the NCICB caGrid 1.0 QA index service. To change the index service URL, open the `file deploy.properties` at `<CANANOLAB_GRID_SOURCE>` and edit the property `index.service.url` to be your choice of caGrid index server.
3	Edit the file `serviceMetadata.xml` at `<CANANOLAB_GRID_SOURCE>/etc/` as follows: • Replace the contents in the tag `<ns1:hostingResearchCenter>` with your hosting center information. **Note**: The value for the attribute `displayName` is used in the caNanoLab web application as the grid node host name to distinguish different remote caNanoLab grid services.

4	Execute the Ant build file (`build.xml`) located at `<CANANOLAB_GRID_SOURCE>` with the following two build properties and the task `createTomcatDeploymentTar`:
	
	`cananolab.sdk.host =<SERVER_URL> (e.g. localhost:8080)`
	
	`webapp.name = wsrf-canano`
	
	Execution of the Ant build script generates a file `CaNanoLabSvc.tar` at the `<CANANOLAB_GRID_SOURCE>/output` directory.

5	Place the file `CaNanoLabSvc.tar` to the directory `<CATALINA_HOME>` and extract the tar file within the same directory.

6	For the caNanoLab grid service to successfully register with the index server, edit the file `server-config.wsdd` at `<CATALINA_HOME>/webapps/wsrf-canano/WEB-INF/etc/globus_wsrf_core as follows`:
	
	• Add a `logicalHost` parameter, with the value being the public IP address or the public hostname for the Tomcat server. For example:
	
	`<globalConfiguration>`
	`...`
	`<parameter name="sendXsiTypes" value="true"/>`
	
	`<parameter name="logicalHost"`
	`value="<TOMCAT_SERVER_URL>"/>`
	
	`<parameter name="publishHostName" value="true"/>`

Step	Action
7	Start the tomcat server to deploy the grid service.

Verification

Open the URL http://<TOMCAT_SERVER_URL>/wsrf-canano/services/cagrid/CaNanoLabSvc and you should see a page that says

```
cagrid/CaNanoLabSvc

Hi there, this is an AXIS service!
```

Log into the caNanoLab application, and select **REMOTE SEARCH > Nanoparticles**. You should see the display name for the research center you specified earlier listed as one of the remote grid node hosts.

For more information and help on caGrid service configuration and deployment, refer to the caGrid wiki page: http://www.cagrid.org.

If you need further assistance on setting a caNanoLab grid service, contact NCICB Application Support at ncicb@pop.nci.nih.gov.

Contacting Application Support

NCICB Application Support

http://ncicb.nci.nih.gov/NCICB/support
Telephone: 301-451-4384
Toll free: 888-478-4423