caARRAY 2.0 DATA PORTAL

Local Installation Guide

NATIONAL® CANCER INSTITUTE

Center for Bioinformatics

Revised January 28, 2008

Table of Contents

Introduction

This caArray 2.0 installation Guide provides you with the instructions to install and configure the caArray 2.0 application in your environment. The caArray installation installs and configures three JBoss application servers, a grid service and creates a database on a preinstalled MySQL server.

NOTE 	Published caArray development documentation can be found on the caArray page of the NCICB web site: http://caarray.nci.nih.gov/

Overview of caArray Installation

The process for installing caArray includes the following tasks described in this document:

- Downloading and installing required software
- Setting environment variables
- Downloading and installing the Universal Provisioning Tool (UPT)
- Downloading caArray 2.0 files
- Configuring JBoss and MySQL as Services
- Advertising the grid service

Before You Proceed 	• Even if you have had previous versions (1.x) of caArray, you must proceed through the pages and steps outlined in this installation guide as if it is a first-time install. It is not possible to “upgrade” an existing installation to caArray 2.0. • Please contact us directly for 1.x data migration support: Web: http://ncicb.nci.nih.gov/NCICB/support/caarraysupport E-mail: ncicb@pop.nci.nih.gov Telephone: 301-451-4384 Toll free: 888-478-4423 • Directions are given in this document for both Linux and Windows operating systems.

caArray 2.0 Software and Technology Requirements

Tested Environment

The caArray 2.0 installation has been tested on Linux Red Hat Enterprise Linux AS 4 64/32-bit (for AMD chipset) and the Windows XP/2003 environments. While the installation may work in other Linux and Windows environments, it has only been tested in these environments.

Required Software—Not Included in caArray

Many of the servers and services that make up caArray 2.0 are automatically installed as part of this installation. However, certain tools that you must manually install and configure are listed in Table 1. The software name, version, description, and URL hyperlinks (for download) are indicated in the table.

You must install the following tools/versions prior to the caArray 2 installation, in the order they are listed. Follow the directions on the corresponding websites for download and installation.

Required Software Name ***Version*** ***URL to Download***	***Description***
Java 2 Platform Standard Edition 5.0 Update 10 (J2SE 5.0) http://java.sun.com/products/archive/j2se/5.0_10/index.html	The J2SE Development Kit (JDK) supports creating J2SE applications.
Apache Ant, 1.7.0 http://ant.apache.org/bindownload.cgi	Apache Ant is a Java-based build tool.
MySQL, 5.0.27 http://dev.mysql.com/downloads/mysql/5.0.html#downloads	MySQL is an open-source database software application.

Table 1 Required Software

Java SDK Installation

To install the Java SDK, follow these steps, proceeding for your operating system as described:

Linux

In Linux, follow these steps:

Step	Action
1	Download the Java SDK from http://java.sun.com/products/archive/j2se/5.0_10/. **Note:** Be sure to download the correct Java SDK for your operating environment. For example, for AMD 64, you would download **`jdk-1_5_0_10-linux-amd64-rpm.bin`**.
2	Login as the root user. **Note:** To install the Java SDK in a system-wide location such as `/usr/local`, you must login as the root user to gain the necessary permissions. If you do not have root access, install the Java SDK in your home directory or a subdirectory for which you have write permissions.
3	When prompted, enter the *root password.*
4	Change to the directory where you want to install Java. For example, enter: `cd /usr/java`
5	Change the permission of the file you downloaded to be executable. For example, enter ***`chmod +x jdk-1_5_0_10-linux-amd64-rpm.bin`***. **Note:** The name of the file will depend upon your target Linux operating system and associated chipset.
6	From the command prompt, enter ***`./jdk-1_5_0_10-linux-amd64-rpm.bin`*** to unzip the file.
7	Execute the rpm installer by entering: ***`rpm -iv jdk-1_5_0_10-linux-amd64.rpm`*** **Note:** The name of the file will depend upon your target Linux operating system and associated chipset.
8	Read through the license and enter **Yes** to proceed with the installation.

Windows

In Windows, follow these steps:

Step	Action
1	Download the Java SDK from http://java.sun.com/products/archive/j2se/5.0_10/. **Note:** Be sure to download the correct Java SDK for your operating environment. For example, for AMD 64, you would download `jdk-1_5_0_10-linux-amd64-rpm.bin`.
2	Login as a Windows Administrator.
3	Run the Java SDK installer for Windows During the installation process, you will be prompted to enter the directory where you wish to install Java. This directory will be used when Setting Environment Variables.

Apache Ant Installation

NOTE 	Apache Ant, version 1.7.0, is the required build tool to install the caArray 2.0 application and services.

To download and extract the Ant build tool, follow these steps:

Step	Action
1	Download Apache Ant from https://gforge.nci.nih.gov/svnroot/lsd/trunk/tools/apache-ant-1.7.0-bin.zip to a directory where you wish to install the tool. *Example*: `/usr/java`
2	Open a command prompt from the location to which you downloaded the `apache-ant-1.7.0-bin.zip` file and enter ***unzip -q apache-ant-1.7.0-bin.zip***. After extracting the zip, you must set the environment variable, described in the following section, so that Ant is available in the system PATH.

Setting the Environment Variables

NOTE 	The purpose of setting operating system environment variables is so that the Java SDK and Ant build tool are available to run from anywhere in the system.

Linux

To set the environmental variables in Linux, follow these steps:

NOTE 	The `JAVA_HOME`, `ANT_HOME` and `PATH` environment variables are set in `/etc/profile` and `/root/.bash_profile` . You may need to create the variables, or modify them if they already exist.

Step	Action

1	As the root user, enter the following in the `/etc/profile` file. A PATH variable should already be created in this file, so be sure to define the JAVA_HOME and ANT_HOME export before the PATH export. Replace *<installation_location>* with the directory where you have installed the Java SDK and Ant
	
	(The location example in the Apache Ant installation (p. 5) is `/usr/java`.)
	
	`export JAVA_HOME=<`*installation_location*`>/jdk1.5.0_10`
	
	`export ANT_HOME==<`*installation_location*`>/apache-ant-1.7.0`
	
	`export PATH=$JAVA_HOME/bin:$ANT_HOME/bin:$PATH`

2	Add the same data from step 1 to the */root*`/.bash_profile` file.
3	Log out and log back in so that the system recognizes your changes.

Verifying the Environment Variables in Linux

To verify that environment variables have been set correctly, follow these steps:

Step	Action
1	From the command line, enter: ***`echo $JAVA_HOME`*** ***`echo $ANT_HOME`*** Both of these commands should return the location where you installed these tools.
2	To verify your Java SDK installation, type ***`java -version`*** from a command prompt. You should see `java version "1.5.0_10".`
3	To verify your Ant installation. type ***`ant -version`*** from a command prompt. You should see: `Apache Ant version 1.7.0 compiled on December 13 2006`

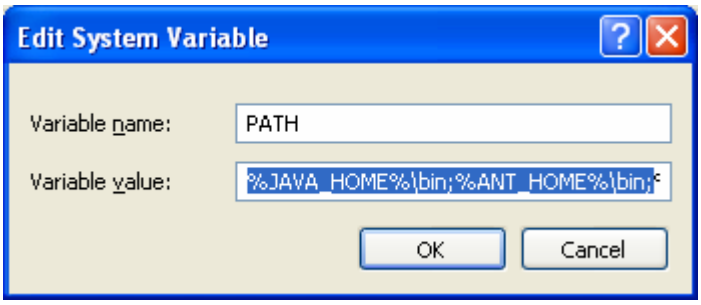
Windows

To set the environmental variables in Linux, follow these steps:

NOTE	
	The JAVA_HOME, ANT_HOME and PATH environment variables are set in the Systems Properties.

:

Step	Action
1	In Windows, select **Control Panel**, then select the **Systems** application. In the Systems window, select the **Advanced** tab.
2	On the **Advanced** tab, click the **Environment Variables** button; to add a new system variable, select the **New** button. In the **Variable name** text box, enter `JAVA_HOME`. In the **Variable value** text box, enter the location of your Java installation.
3	Click the **New** button again. In the **Variable name** text box, enter ***`ANT_HOME.`*** In the **Variable value** text box, enter the location of your Java installation.

Step	Action
4	Select the PATH system environment variable. and select the **Edit** button. This opens the Edit System Variable dialog box, displayed here as an example. 
5	In the Variable value text box, enter: *%JAVA_HOME%\bin;%ANT_HOME%\bin;*

Verifying the Environment Variables in Windows

To verify the environment variables have been set correctly, follow these steps:

Step	Action
1	From the command line, enter: ***echo %JAVA_HOME%*** ***echo %ANT_HOME%*** Both of these commands should return the location where you installed these tools.
2	To verify your Java SDK installation, type ***java -version*** from a command prompt. You should see `java version "1.5.0_10".`
3	To verify your Ant installation, type ***ant -version*** from a command prompt. You should see: `Apache Ant version 1.7.0 compiled on December 13 2006.`

MySQL Installation

NOTE	As part of the installation process, the default character set is set to `latin1` for the caArray MySQL database.

Linux

Use RPM to install MySQL server - version 5.0.27.

Configuration

After MySQL has been installed, it must be configured using the following steps:

Step	Action
1	*Lowercase Table Names in MySQL* ***sudo vi /etc/init.d/mysqld*** a. Locate the `start()` section and modify the `mysqld_safe` command: ***/usr/bin/mysqld_safe --lower_case_table_names=1 ...*** b. Restart the MySQL service for the changes to take effect: `sudo /etc/init.d/mysqld restart`
2	*Increase the Java Memory Allocation* Open the `/etc/my.cnf` and add the following: ***[mysqldump]*** ***max_allowed_packet=64M***

Step	Action
3	*Update the Help Desk Contact Information (SQL)* The Help Desk information provided by default in the caArray database must be changed for your environment. To do this, connect to your MySQL server and the caArray database and run the following script (replacing `email_address` with a return email address that is accessible at your location). ***update config_parameter set raw_value = '[email_address]' where param = 'REG_EMAIL_TO';*** ***update config_parameter set raw_value = '[email_address]' where param = 'EMAIL_FROM';***

Windows

Use the MySQL 5.0.27 installer for Windows.

Configuration

Step	Action
1	*USE LOWERCASE TABLE NAMES IN MYSQL* a. Locate the `start()` section and modify the mysqld_safe command: ***/usr/bin/mysqld_safe --lower_case_table_names=1 ...*** b. Restart the MySQL Windows service for the changes to take effect. Use **Administrator Tools|Services.**
2	*INCREASE THE JAVA MEMORY ALLOCATION* a. Open the [*MySQL installation directory*]/my.ini file and add the following: ***[mysqldump]*** ***max_allowed_packet=64M*** b. Restart the MySQL Windows service for the changes to take effect. Use **Administrator Tools|Services**

Step	Action
3	*UPDATE THE HELP DESK CONTACT INFORMATION (SQL)* The Help Desk information provided by default in the caArray database must be changed for your environment. To do this, connect to your MySQL server and the caArray database and run the following script (replacing `email_address` with a return email address that is accessible at your location). ***update config_parameter set raw_value = '[email_address]' where param = 'REG_EMAIL_TO';*** ***update config_parameter set raw_value = '[email_address]' where param = 'EMAIL_FROM';***

Installing caArray 2.0 Application and Services

To install the caArray 2.0 application and services, follow the steps in this section::

- Download and install the Universal Provisioning Tool 3.2 (UPT)
- Download caArray 2.0 files from GForge
- Configure JBoss and MySQL servers
- Advertise the Grid Service

Downloading and Installing the UPT files

To download and install the **UPT 3.2** files, follow these steps:

Step	Action
1	The installation file for UPT 3.2 is over 15MB. Download this file from GForge: https://gforge.nci.nih.gov/svnroot/lsd/trunk/dist/upt-installer.zip. Remember the download location as you will be using this file to run the installation in the steps that follow. This location will later be referred to as the *installation_directory*.
2	From the directory where you downloaded the `upt-installer.zip` file, open a command prompt and enter ***./install*** (***install.sh*** in Linux and ***install.bat*** [without the `./`] in Windows). The installation unzips the file and places it in the UPT directory underneath the installation directory.
3	If you need to modify the default properties, open the *<installation_directory>*`/upt/upt.properties` file and modify the values for your environment and save the file.

Step	Action
4	Navigate to `<installation_directory>/upt` and type ***`ant`***. This will run the installation.

Downloading caArray 2.0 files

To download the caArray 2.0 files, follow these steps:

Step	Action
1	The installation file for caArray 2.0 is over 200MB. Download this file from GForge: https://gforge.nci.nih.gov/svnroot/lsd/trunk/dist/caarray2-installer.zip. Remember the download location as you will be using this file to run the installation in the steps that follow.
2	These server components are installed and configured as part of the caArray2 installation. You do not need to do anything further to download or install these components. • JBoss 4.0.4 (hosts the caArray grid service) • JBoss 4.0.5 (hosts the caArray application) • JEMS installer 1.0.2 GA – supports EJB 3.0 specification

Installing caArray 2.0

To install caArray 2.0, follow these steps:

Step	Action
1	From the directory where you downloaded the `caarray2-installer.zip` from Downloading caArray 2.0 files, open a command prompt and enter ***`./install`*** (***`install.sh`*** in Linux and ***`install.bat`*** [without the `./`] in Windows). The installation unzips the `caarray2-installer.zip` file and places it in a directory called `caarray` which is one directory below the installation directory.

Step	Action

2	To modify the default properties, open the `<installation_directory>/caarray2/caarray2.properties` file, modify the values for your environment and save the file. At a minimum, you will need to modify the following values based on your MySQL server environment:
	• `database.system.user=REPLACE_THIS_VALUE`
	  ○ This value should correspond to a MySQL username that has been given full database privileges. For example, `caarrayadmin` may be a good name for this user.
	• `database.system.password`
	  ○ This value must correspond to the password for the `database.system.user` user.
	• `database.server`
	  ○ This value must correspond to the domain name of machine that hosts the MySQL server.
	• `database.port`
	  ○ This value must correspond to the port for the `database.server`. 3306 is the default port, but check with your database administrator to be certain.
	Note: You shouldn't need to modify the other defaults values as we have chosen unique ports to reduce the risk of other applications using the same values

3	From the `<installation_directory>/caarray2` directory, enter ***ant*** from the command prompt. This initiates the installation process. The anticipated duration is anywhere 1-15 minutes depending on your system's speed, power and memory. The installer will create a caArray database on your MySQL server, start and configure two JBoss servers and start up a grid service for the caArray application. To access caArray, open your web browser to http://<jboss.server.hostname>:<jboss.server.port>/caarray. **Note:** `jboss.server.hostname` and `jboss.server.port` are values in the `<installation_directory>/caarray2/caarray2.properties` file.
4	To verify that the application was installed correctly, enter ***caarrayadmin*** as the user and ***caArray2***! as the password.

Configuring JBoss

NOTE	
	For optimal performance, you must modify your JBoss 4.0.5 configuration to increase the amount of available memory for the caArray application. Directions for doing this are in the following step 1.

To configure JBoss, follow these steps:

Step	Action
1	Add the following entry to the JBoss run.conf file. For example, the run.conf will be located at `<installation_directory>/caarray3-app/jboss-4.0.5.GA/bin/run.conf`. `JAVA_OPTS="-Xms512m -Xmx2048m –XX:ThreadStackSize=128 -XX:SurvivorRatio=10 -XX:PermSize=128m -XX:MaxPermSize=128m -Dsun.rmi.dgc.client.gcInterval=3600000 -Dsun.rmi.dgc.server.gcInterval=3600000"`
2	Restart your JBoss 4.0.5 server for the changes to take effect.

Configuring JBoss and MySQL to run as services

NOTE 	Both MySQL and the three JBoss servers that make up caArray must run continually as services. The instructions in this section cover all of these scenarios. For caArray 2.0, there are a total of four servers: • JBoss for UPT • JBoss 4.0.4 • JBoss 4.0.5 • MySQL 5.0.27

Running JBoss as a service

NOTE 	The default caArray installation runs JBoss as a command line process using the user currently logged on. Therefore, when you log out as this user, JBoss will no longer be available for caArray. For that reason, it is recommended that you configure your JBoss servers to run as a Linux or Windows service. The instructions are contained in this section.

To run JBoss as a service, follow these steps:

Step	Action
1	**Linux** See http://wiki.jboss.org/wiki/Wiki.jsp?page=StartJBossOnBootWithLinux.
2	**Windows** To run an existing JBoss command line installation as a service, open a command prompt and enter: ***sc create caarray2-jboss binpath= "[installation_home]/jboss-4.0.5/bin/run.bat" --service displayname= "Caarray2 JBoss Server" depend=Tcpip start=auto*** **Note:** You need to modify *installation_home* to the location of your installation:

Running MySQL as a service

NOTE 	It is assumed that your MySQL server was installed as a service. If it was not, follow these recommendations for installing this server as a service.

To run JBoss as a service, follow these steps:

Step	Action
1	**Linux** See http://www.redhat.com/docs/manuals/enterprise/RHEL-AS-2.1-Manual/cluster-manager/s1-service-mysql.html.
2	**Windows** When installing MySQL server on Windows, choose the option to run MySQL as a Windows service.

Post-Installation: Advertising the Grid Service

To advertise the caArray grid service, you must update the `serviceMetadata.xml` file, and then restart the JBoss 4.0.4 server instance.

Step	Action
1	Edit `${caarray2.home}/jboss-4.0.4.GA/etc/serviceMetadata.xml`. The sections to update are **service description**, **point of contact** and **service context description**.
2	After making these changes, restart the JBoss server (which hosts the grid service).

Contacting Application Support

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