CAARRAY 2.0 DATA PORTAL

Local Installation Guide



Revised February 26, 2008

Table of Contents

INTRODUCTION	
CAARRAY 2.0 SOFTWARE AND TECHNOLOGY REQUIREMENTS	2
Java SDK Installation	
Linux Windows	3
Apache Ant Installation	4
SETTING THE ENVIRONMENT VARIABLES	5
Linux	
Windows	6
MYSQL INSTALLATION	8
INSTALLING CAARRAY 2.0 APPLICATION AND SERVICES	8
Downloading and Installing the UPT files	
Downloading caArray 2.0 files	10
Installing caArray 2.0	11
caArray Port Usage	
JBoss Errors During Installation	
Configuring MySQL	
Configuring JBoss and MySQL to Run as Services	10
Post-Installation: Advertising the Grid Service	17
CONTACTING APPLICATION SUPPORT	19

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
- · 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: 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	Description	
URL to Download		
Java 2 Platform Standard Edition 5.0 Update 10 (J2SE 5.0)	The J2SE Development Kit	
http://java.sun.com/products/archive/j2se/5.0_10/index.html	(JDK) supports creating	
	J2SE applications.	
Apache Ant, 1.7.0	Apache Ant is a Java-	
http://ant.apache.org/bindownload.cgi	based build tool.	
MySQL, 5.0.27	MySQL is an open-source	
http://downloads.mysql.com/archives.php?p=mysql-5.0&v=5.0.27	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.
	Login as the root user (or sudo).
2	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
	Change the permission of the file you downloaded to be executable.
5	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	Unzip the Java SDK bin file. For example, from the command prompt, enter ./jdk-1_5_0_10-1inux-amd64-rpm.bin to unzip the file.
	Execute the rpm installer. For example, enter: rpm -iv jdk-1_5_0_10-1inux-amd64.rpm
7	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 Windows, you might download jdk-1_5_0_10-windows-i586-p.exe.
2	Login as a Windows Administrator.
	Run the Java SDK installer for Windows
3	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 the Environment Variables

Apache Ant Installation



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 <i>unzip -q apache-ant-1.7.0-bin.zip</i> . 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 /<username>/.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_directory> with the directory where you have installed the Java SDK and Ant.</installation_directory>
	(The location example in the Apache Ant installation (p. 5) is /usr/java.)
	export JAVA_HOME= <installation_directory>/jdk1.5.0_10</installation_directory>
	export ANT_HOME= <installation_directory>/apache-ant-1.7.0</installation_directory>
	export PATH=\$JAVA_HOME/bin:\$ANT_HOME/bin:\$PATH
2	Add the same data from step 1 to the / <username>/.bash_profile file.</username>
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
	From the command line, enter:
1	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, enter <code>java -version</code> from a command prompt. You should see <code>java version "1.5.0_10"</code> .
3	To verify your Ant installation, enter: ant -version from a command prompt. You should see: Apache Ant version 1.7.0 compiled on December 13 2006.

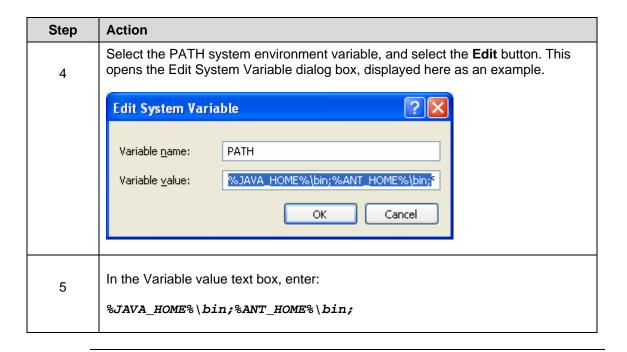
Windows

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



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.
	a. In the Variable <u>n</u> ame text box, enter JAVA_HOME.
	b. In the Variable value text box, enter the location of your Java installation.
2	Click the New button again.
3	a. In the Variable name text box, enter ANT_HOME.
	b. In the Variable value text box, enter the location of your Java installation.



Verifying the Environment Variables in Windows

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

Step	Action
	From the command line, enter:
1	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, enter <code>java -version</code> from a command prompt. You should see <code>java version "1.5.0_10"</code> .
	To verify your Ant installation, enter ant -version from a command prompt.
3	You should see: Apache Ant version 1.7.0 compiled on December 13 2006.

MySQL Installation

A MySQL 5.0.27 server must be installed and running in order for the installation to work successfully. As part of the installation process, there are several built-in pre-install checks that are performed to ensure you have a valid server and connection information.

NOTES



- As part of the installation process, the default character set is set to latin1 for the caArray MySQL database.
- MySQL must be configured, but not until after caArray is installed.
 See Configuring MySQL on page 14.

Installing caArray 2.0 Application and Services

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

- Downloading and Installing the UPT files
- Download caArray 2.0 files from GForge
- Configuring MySQL
- Configuring JBoss and MySQL to Run as Services
- Post-Installation: Advertising the Grid Service

Downloading and Installing the UPT files

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

Step	Action
	The installation file for UPT 3.2 is over 30MB.
1	From the https://gforge.nci.nih.gov/svnroot/lsd/trunk/dist/ directory in GForge, download the upt_distribution_[version].zip file.
	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>.</installation_directory>
2	From the directory where you downloaded the upt_distribution_[version].zip file, open a command prompt extract this file to a temporary location. For example, you may enter a command such as unzip -q upt_distribution_[version].zip (you must have a ZIP tool installed). This location will be referred to as the <installation_directory> henceforth.</installation_directory>
	Note : You must give the proper permissions to the upt_distribution_[version].zip file in order to extract it (for example, chmod in Linux).

Step	Action	
3	If you need to modify the default properties, open the <pre><installation_directory>/upt/upt-install.properties file and modify the values for your environment and save the file. At a minimum, you will need to modify the following values:</installation_directory></pre>	
	• upt.home	
	o The location to which you are going install UPT (your <application_root_directory>). For example, in Windows, the <application_root_directory> can be C:/apps/upt. Linux users can use /apps/upt or any other folder to which you have write permissions.</application_root_directory></application_root_directory>	
	<pre>Important: <application_root_directory> must be different than <installation_directory> or the installation will fail.</installation_directory></application_root_directory></pre>	
	database.system.user	
	 This value should correspond to a MySQL username that has been given full database privileges. For example, uptadmin may be a good name for this user. 	
	database.system.password	
	 This value <u>must</u> correspond to the password for the database.system.user user. 	
	• database.server	
	 This value <u>must</u> correspond to the domain name of machine that hosts the MySQL server. 	
	database.port	
	 This value <u>must</u> correspond to the port for the database.server. 3306 is the default port, but check with your database administrator to be certain. 	
	• database.name	
	 Choose a name for the UPT MySQL database. 	
	• database.user	
	 Choose a username to access database.name. This can be any valid name that you choose, but it must be different than database.system.user. 	
	database.password	
	 Choose a password to access database.name for the username identified in database.user 	
	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. However, be sure to check the upt-install.properties to verify that the ports in this file are not being used by other applications, otherwise you will experience problems.	

Step	Action
4	Navigate to <installation_directory>/upt and type ant. This will run the installation. You will verify UPT installation after Installing caArray 2.0.</installation_directory>
5	After successfully installing UPT, make a backup of <pre><installation_directory>/upt/upt-install.properties for future reference.</installation_directory></pre>

UPT Port Usage

NOTE



Verify that default port values defined in upt-install.properties files are not in use on your system by running netstat -a from the command line. If the ports are in use prior to installation, you will likely experience problems with your installation.

Downloading caArray 2.0 files

To download the caArray 2.0 files, follow these steps:

Step	Action
	The installation file for caArray 2.0 is over 300MB.
1	From the https://gforge.nci.nih.gov/svnroot/lsd/trunk/dist/directory in GForge, download the caarray_distribution_[version].zip file.
	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 caarray_distribution_2_0_0.zip from Downloading caArray 2.0 files, open a command prompt extract this file to a temporary location. For example, you may enter a command such as caarray_distribution_2_0_0.zip (you must have a ZIP tool installed). For example, use unzip - q caarray_distribution_2_0_0.zip. This location will be referred to as the <installation_directory> henceforth. Once you unzip this file, it creates a directory called caarray which is a directory below the <installation_directory></installation_directory></installation_directory>
2	To modify the default properties, open the <installation_directory>/caarray/caarray2-install.properties file, modify the values for your environment and save the file. At a minimum, you will need to modify the following values:</installation_directory>
	• caarray2.home
	o The location to which you are going to install caArray (your <application_root_directory>). For example in Windows, the <application_root_directory> can be C:/apps/caarray2. Linux users can use /apps/caarray2 or any other folder to which you have write permissions.</application_root_directory></application_root_directory>
	<pre>Important: <application_root_directory> must be different than <installation_directory> or the installation will fail.</installation_directory></application_root_directory></pre>
	• database.system.user
	 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
	o This value <u>must</u> correspond to the password for the database.system.user user.
	• database.server
	 This value <u>must</u> correspond to the domain name of machine that hosts the MySQL server.
	• database.port
	o This value <u>must</u> correspond to the port for the database.server. 3306 is the default port, but check with your database administrator to be certain.
	• database.name
	 Choose a name for the caArray MySQL database.
	• database.user

Step	Action
	 Choose a username to access database.name. This can be any valid name that you choose, but it must be different than database.system.user.
	database.password
	 Choose a password to access database.name for the username identified in database.user
	• mail.smtp.host
	 This value <u>must</u> correspond to an SMTP server available in your network.
	• mail.smtp.port
	 This value <u>must</u> correspond to the SMTP server. The default is 25, but this may be different in your environment.
	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. However, be sure to check the <code>caarray2-install.properties</code> to verify that the ports in this file are not being used by other applications.
3	From the <installation_directory>/caarray 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.</installation_directory>
	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 verify UPT installation, go to: <a href="http://<jboss.server.hostname>:<jboss.server.port>/upt">http://<jboss.server.hostname>:<jboss.server.port>/upt</jboss.server.port></jboss.server.hostname> . To access caArray, open your web browser to

caArray Port Usage

NOTE



Verify that default port values defined in caarray2-install.properties files are not in use on your system by running netstat -a from the command line. If the ports are in use prior to installation, you will likely experience problems with your installation.

JBoss Errors During Installation

NOTE



You may receive an error such as Exception in thread "main" java.lang.NoClassDefFoundError: org/jboss/Shutdown. This should not be a problem as the installer attempts to remove previously installed servers to prevent problems during the installation. If this is your first time installing caArray, you may receive and disregard this error message.

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 in **Windows**, follow these steps:

Step	Action
1	Add the following entry to the JBoss run.bat file. For example, the run.conf will be located at <installation_directory>/jboss-4.0.5.GA/bin/run.bat.</installation_directory>
	-server -Xms2048m -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 - Djava.awt.headless=true
2	Restart your JBoss 4.0.5 server for the changes to take effect.

Configuring MySQL

Linux

After MySQL has been installed (see page 8), in Linux it must be configured using the following steps:

Step	Action
	Lowercase Table Names in MySQL
1	Edit the /etc/init.d/mysqld file as follows:
	a. Locate the start() section and modify the mysqld_safe command (do not include the ellipses):
	/usr/bin/mysqld_safelower_case_table_names=1
	b. Restart the MySQL service for the changes to take effect:
	Restart /etc/init.d/mysqld
2	Increase the Java Memory Allocation
	Open the $/etc/my.cnf$ and add the following (you may need to create this file, if it hasn't been created):
	[mysqldump] max_allowed_packet=64M [mysqld] max_allowed_packet=64M
	[mysq1]
	max_allowed_packet=64M
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).
	<pre>update config_parameter set raw_value = '[email_address]' where param = 'REG_EMAIL_TO';</pre>
	<pre>update config_parameter set raw_value = '[email_address]' where param = 'EMAIL_FROM';</pre>

Windows

After MySQL has been installed (see page 8), in Windows it must be configured using the following steps:

Step	Action
	USE LOWERCASE TABLE NAMES IN MYSQL
1	a. Locate the start() section and modify the mysqld_safe command:
	/usr/bin/mysqld_safelower_case_table_names=1
	b. Restart the MySQL Windows service for the changes to take effect. Use Administrator Tools Services.
	INCREASE THE JAVA MEMORY ALLOCATION
2	<pre>a. Open the [MySQL installation directory]/my.ini file and add the following: [mysqldump] max_allowed_packet=64M [mysqld] max_allowed_packet=64M [mysql] max_allowed_packet=64M</pre>
	b. Restart the MySQL Windows service for the changes to take effect. Use Administrator Tools Services
	UPDATE THE HELP DESK CONTACT INFORMATION (SQL)
3	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).
	<pre>update config_parameter set raw_value = '[email_address]' where param = 'REG_EMAIL_TO';</pre>
	<pre>update config_parameter set raw_value = '[email_address]' where param = 'EMAIL_FROM';</pre>

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 (for Grid services)
- JBoss 4.0.5 (for caArray application)
- 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 .
	Windows
2	To run an existing JBoss command line installation as a service, follow the directions for creating a user-defined service at http://support.microsoft.com/kb/137890/EN-US/
	Note: You need to have access to the Windows Resource Kit.

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 MySQL as a service, follow these steps:

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

Post-Installation: Advertising the Grid Service

To advertise your caArray grid service, you must take the following steps, then restart the JBoss 4.0.4 server instance. You can check if your grid service is advertised correctly at http://cagrid-portal.nci.nih.gov/web/guest/home.

Step	Action
	web.xml
1	The web.xml file contains the port and protocol your grid service will be advertised as. Change this file if your service is on a port other than the default port of 18080 and/or the default protocol of http
	<init-param></init-param>
	<pre><param-name>defaultProtcol</param-name></pre>
	<pre><param-value>http</param-value></pre>
	<init-param></init-param>
	<pre><param-name>defaultPort</param-name></pre>
	<pre><param-value>18080</param-value></pre>
	File Locations:
	\${caarray2.home}/jboss- 4.0.4.GA/server/default/deploy/wsrf.war/WEB-INF
	server-config.wsdd
2	Make sure your container is publishing the right host name. Your service must register with a publicly accessible address or DNS-resolvable host name, so the Index Service (and other clients) can connect to it. Add the following lines to this file if you want your service to have a specific name or if your service is trying to register a private IP address which is not allowed. You should see errors in your JBoss 4.0.4 log if you are trying to register a private IP address.
	<pre><parameter name="logicalHost" value="somehost.cagrid.org"></parameter></pre>
	<pre><parameter name="publishHostName" value="true"></parameter></pre>
	File Location:
	<pre>\${caarray2.home}/jboss- 4.0.4.GA/server/default/deploy/wsrf.war/WEB- INF/etc/globus_wsrf_core</pre>

Step	Action
	serviceMetadata.xml
3	This file contains the service's contact information. The two sections to update are at the top and bottom of the file.
	Top of file under <ns2:pointofcontactcollection>:</ns2:pointofcontactcollection>
	<pre><ns3:pointofcontact affiliation="" email="" firstname="" lastname="" phonenumber="" role="" xmlns:ns3="gme://caGrid.caBIG/1.0/gov.nih.nci.cagrid.metadata.common"></ns3:pointofcontact></pre>
	Bottom of file under <ns1:hostingresearchcenter>:</ns1:hostingresearchcenter>
	<ns14:pointofcontact affiliation="" email="" firstname=" " lastname=" " phonenumber="" role=" "></ns14:pointofcontact>
	File Location:
	\${caarray2.home}/jboss- 4.0.4.GA/server/default/deploy/wsrf.war/WEB- INF/etc/caGrid_CaArraySvc
	CaArraySvc_registration.xml
4	This file has the URL to the Index Server where you can see if your service is advertised. Ensure this file contains the following index server:
	<pre><wsa:address>http://cagrid- index.nci.nih.gov:8080/wsrf/services/DefaultIndexService</wsa:address></pre>
	File Location:
	<pre>\${caarray2.home}/jboss- 4.0.4.GA/server/default/deploy/wsrf.war/WEB- INF/etc/cagrid_CaArraySvc/</pre>
5	After making these changes, restart the JBoss 4.0.4 server (this hosts the grid service).
	For more troubleshooting information : http://www.cagrid.org/mwiki/index.php?title=CaGrid:How-To:TroubleshootIndexService

Contacting Application Support

NCICB http://ncicb.nci.nih.gov/NCICB/support

Application Telephone: 301-451-4384 Toll free: 888-478-4423