

caArray 006 - Step by Step Guide to Install caArray 2.x Grid Service (caArray 2.2 and Under)

Unable to render {include} The included page could not be found.

Step by Step Guide to Install caArray 2.x Grid Service (caArray 2.2 and Under)

Topic: caArray Installation and Upgrade

Release: Up to caArray 2.2

Date entered: 01/26/2009

Contents of this Page

- [Introduction](#)
- [Prerequisite](#)
- [Abbreviation Used](#)
- [Step 1. Make appropriate changes to install.properties file](#)
- [Step 2. Run ant to compile source code](#)
- [Step 3. Make appropriate changes to web.xml file](#)
- [Step 4. Make appropriate changes to server-config.wsdd](#)
- [Step 5. Make appropriate changes to CaArraySvc_registration.xml](#)
- [Step 6. Make appropriate changes to serviceMetadata.xml](#)
- [Step 7. Restart JBoss server for grid service](#)
- [Step 8. Testing caGrid Service Installed](#)
- [Have a comment?](#)

Introduction

There are three components installed during caArray's installation process: UPT, caArray and caArray Grid Service (refer to [caArray 001 - UPT required for caArray or other application tools in LSD Bundle](#)). UPT is a mandatory application used for caArray's user access control. The caGrid Service is used to advertise your caArray application on the [NCI caGrid Portal](#). The caGrid Portal provides a visual display of the services that are running on the caGrid infrastructure and the institutions that are participating in the caBIG program. You can easily find caBIG services, participants, and points of contact on the caGrid Portal. You can also query caGrid data services, share your queries, and search for the queries of other community members. Advertising your caGrid Service on caGrid Portal is optional. You may choose to participate at a later time, or you may choose not to participate at all.

The official **caGrid Service configuration steps** can be found in [caArray Installation Guide \(page 23\)](#) (login required). This knowledge base entry offers a supplemental step by step installation guide based on our hands-on experience here at Columbia University.

Prerequisite

caArray 2.x application must be running prior to the activation of your caGrid service on the caGrid Portal. However, the grid service is not required for the caArray application to run.

Abbreviation Used

The following abbreviations will be used through the text:

Abbreviation	Comment
<installation_directory>	The location where you keep the files downloaded. Example: /app_data1/caArray2.1
<application_base_path> = \${application.base.path}	The location where you are going to install caArray and caGrid service. Example: /app_data1/caarray_app/caarray
<grid.home> = \${application.base.path}	The location for the caGrid service files, see step 2 It is predefined in install.properties file.

Step 1. Make appropriate changes to install.properties file

File Location: <installation_directory>/install.properties

This is the step to configure Install.properties file before the deployment of caArray Grid service begins. caGrid service and caArray share the same Install.properties file.

It is important to specify following properties:

Property	Comments
application.base.path=/app_data1/caarray_app/caarray	The location where you are going to install caArray (i.e. the deployment of caGrid service and caArray application). It need to be pre-created before installation start. Important: <application_base_path> must be different from <installation_directory> or the installation will fail.
grid.index URL =http://cagrid-index.nci.nih.gov:8080/wsrf/services/DefaultIndexService	NCI caGrid service's production webserver name
domain.name=afapp1.c2b2.columbia.edu	Our private DNS name where the caArray is installed.
grid.static.grid.hostname=localhost	We use localhost here because our caGrid service is on the same machine as caArray (a default configuration). The caGrid service's hostname must be used if it is installed on a different machine from caArray application.
grid.static.grid.port=80	Our publicly accessible port to DMZ Zone

Step 2. Run ant to compile source code

File location: <installation_directory>/ant

During the compiling process, Java class files will be deployed into following two directories, as defined in Install.properties file:

Directory	Comments
grid.home=\${application.base.path}/jboss-4.0.4.GA	For caGrid service (created during the deployment)
jboss.home=\${application.base.path}/jboss-4.0.5.GA	For Jboss service (created during the deployment)

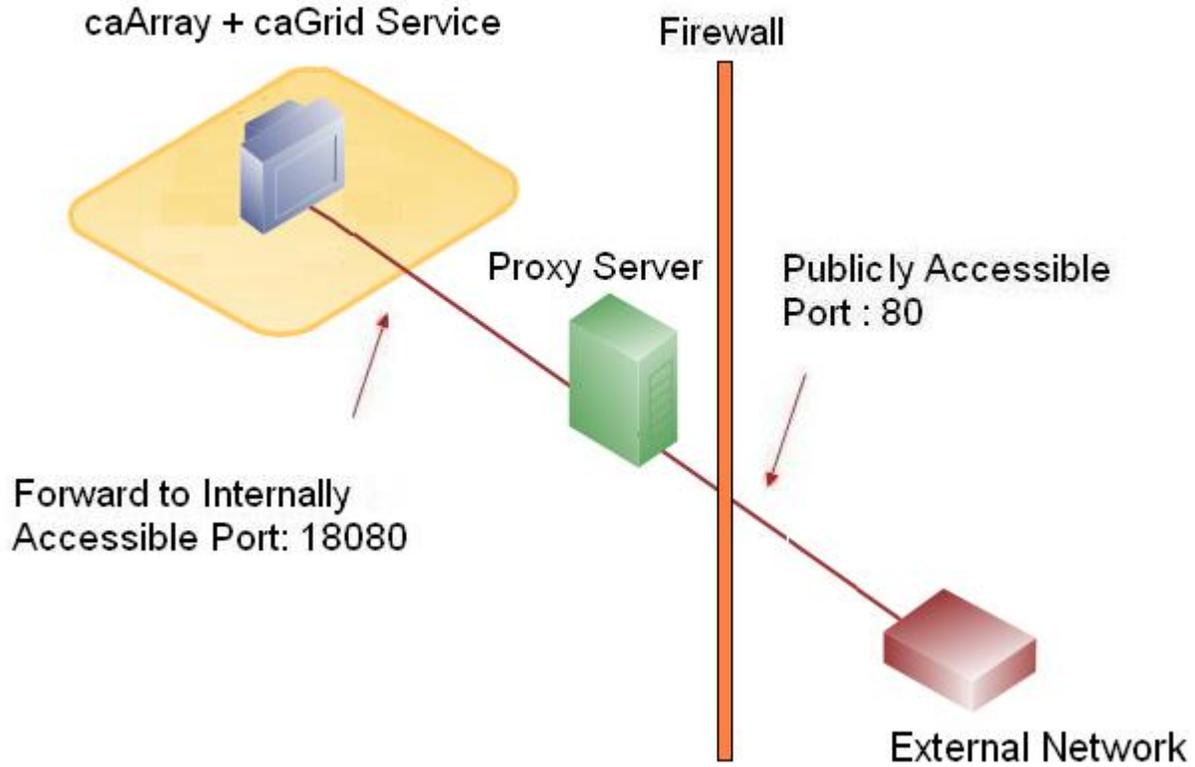
Step 3. Make appropriate changes to web.xml file

File Location: <grid.home>/server/default/deploy/wsrf.war/WEB-INF/web.xml

```
<param-name>defaultPort</param-name>
<param-value>80</param-value>
```

The web.xml file contains the port and protocol your grid service will be advertised as. The port defined here should be the same publicly accessible port as defined in your property file (grid.static.grid.port in **Step 1**). It is port 80 in our case, as defined in our property file.

Figure 1a. caArray Application & Port Forwarding



Our caGrid service server physically resides on an internal server (afapp1.c2b2.columbia.edu:18080) behind the firewall. This address, however, is not reachable from outside of Columbia University's network. Our publicly accessible address (a logical hostname) is:

```
Historical link  
http://caarraygrid.c2b2.columbia.edu:80/wsrf/services/cagrid/CaArraySvc
```

Since our caGrid service is actually running on the internal server's port 18080, a proxy server will forward any incoming request from the publicly accessible port 80 to port 18080 on the internal host.

The port used in web.xml should be the port from the publicly accessible address, not the port on the internal server.

Step 4. Make appropriate changes to server-config.wsdd

File Location: <grid.home>/server/default/deploy/wsrf.war/WEB-INF/etc/globus_wsrf_core/server-config.wsdd

This is to add your caGrid service's publicly accessible DNS name into configuration file.

```
<parameter name="logicalHost" value="caarraygrid.c2b2.columbia.edu"/>
```

Step 5. Make appropriate changes to CaArraySvc_registration.xml

File Location: <grid.home>/server/default/deploy/wsrf.war/WEB-INF/etc/cagrid_CaArraySvc/CaArraySvc_registration.xml

This file has the URL to the Index Server where you can see if your service is advertised. Make sure it contains following:

```
<ServiceGroupEPR>
<wsa:Address>http://cagrid-index.nci.nih.gov:8080/wsrf/services/DefaultIndexService</wsa:Address>
</ServiceGroupEPR>
```

Step 6. Make appropriate changes to serviceMetadata.xml

File Location: <grid.home>/server/default/deploy/wsrf.war/WEB-INF/etc/cagrid_CaArraySvc/serviceMetadata.xml

This file contains the contact information for your caGrid service. Your caGrid Service will not be recognized in the caGrid Portal without the proper identification defined here.

The following is in the top of the file.

```
<ns2:pointOfContactCollection>
<ns3:PointOfContact affiliation="Columbia University Medical Center" email="my_contact@my.organization"
firstName="My FirstName" lastName="My LastName" phoneNumber="My Phone" role="DBA"
xmlns:ns3="gme://caGrid.caBIG/1.0/gov.nih.nci.cagrid.metadata.common"/>
</ns2:pointOfContactCollection>
```

The following is at the bottom of the file.

```
<ns1:hostingResearchCenter>
<ns14:ResearchCenter displayName="Columbia University Medical Center" shortName="CUMC"
xmlns:ns14="gme://caGrid.caBIG/1.0/gov.nih.nci.cagrid.metadata.common">
<ns14:Address country="US" locality="New York" postalCode="10032" stateProvince="NY" street1="My Street1."
street2="" />
<ns14:pointOfContactCollection>
<ns14:PointOfContact affiliation="Columbia University Medical Center" email="my_contact@my.organization"
firstName="My FirstName" lastName="My LastName" phoneNumber="My Phone" role="DBA"/>
</ns14:pointOfContactCollection>
</ns14:ResearchCenter>
</ns1:hostingResearchCenter>
```

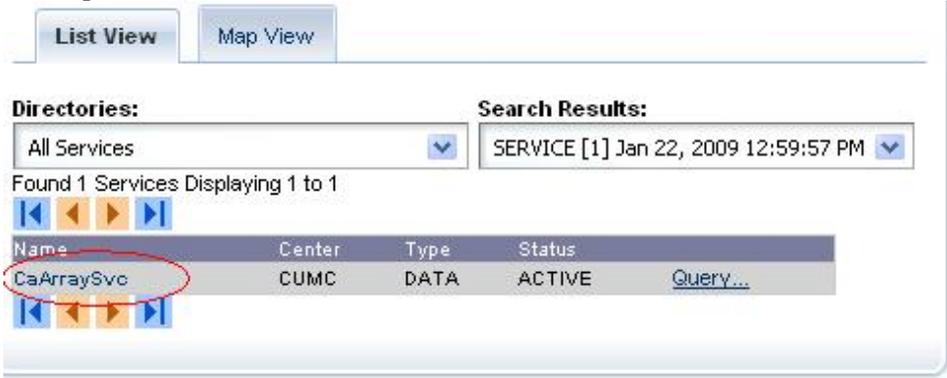
Step 7. Restart JBoss server for grid service

```
cd ${application.base.path}/jboss-4.0.4.GA/bin
nohup ./run.sh &
```

Step 8. Testing caGrid Service Installed

The table below summarizes four methods we used to test if our newly installed caGrid Service (**caarraygrid**) works properly:

Objects	Procedures	Results
Does caarraygrid Registered in NCI Index?	<ul style="list-style-type: none"> export GLOBUS_LOCATION=/app_data1/ws-core-4.0.3 cd /app_data1/ws-core-4.0.3/bin ./wsrf-query -a -z none -s http://cagrid-index.nci.nih.gov:8080/wsrf/services/DefaultIndexService grep 'caarraygrid' 	<p>Following Information should be returned:<ns8:Address xmlns:ns8="http://schemas.xmlsoap.org/ws/2004/03/addressing"></p> <p>http://caarraygrid.c2b2.columbia.edu:80/wsrf/services/cagrid/CaArraySvc</p> <p></ns8:Address></p>

<p>Is our registration accurate with ServiceMetadata?</p>	<ul style="list-style-type: none"> • ./wsrf-get-property -a -z none -s http://caarraygrid.c2b2.columbia.edu:80/wsrf/services/cagrid/CaArraySvc {gme://caGrid.caBIG/1.0/gov.nih.nci.cagrid.metadata} ServiceMetadata 	<p>It should return our contact information, such as:<ns1:ServiceMetadata xmlns:ns1="gme://caGrid.caBIG/1.0/gov.nih.nci.cagrid.metadata"> <ns1:serviceDescription> <ns2:Service description="" name="CaArraySvc" version="1.1" xmlns:ns2="gme://caGrid.caBIG/1.0/gov.nih.nci.cagrid.metadata.service"> <ns2:pointOfContactCollection> <ns3:PointOfContact affiliation="Columbia University Medical Center"</p>										
<p>Is there any data returned with DomainModel?</p>	<ul style="list-style-type: none"> • ./wsrf-get-property -a -z none -s http://caarraygrid.c2b2.columbia.edu:80/wsrf/services/cagrid/CaArraySvc {gme://caGrid.caBIG/1.0/gov.nih.nci.cagrid.metadata.dataservice} DomainModel 	<p>It should return some of our data from caArray, such as: <ns1:DomainModel projectDescription="Version 2.0 caArray Model" projectLongName="caArray" projectShortName="caArray" projectVersion="2" xmlns:ns1="gme://caGrid.caBIG/1.0/gov.nih.nci.cagrid.metadata.dataservice"> <ns1:exposedUMLAssociationCollection> <ns1:UMLAssociation bidirectional="false"> <ns1:targetUMLAssociationEdge> <ns1:UMLAssociationEdge maxCardinality="1" minCardinality="0" roleName="annotation"> <ns1:UMLClassReference refid="5423786E-6C04-59EA-E044-0003BA3F9857"/></p>										
<p>Can I see my caGrid Service online?</p>	<ul style="list-style-type: none"> • Go to http://cagrid-portal.nci.nih.gov/web/guest/home Click the "SERVICES" tab -> "Search" tab -> [MATKC:'Keyword' = "caarraygrid" and 'Search Fields' = "URL"] -> "Search" button 	<p>Checking that caGrid is online</p>  <p>The screenshot shows a web interface with 'List View' and 'Map View' buttons. Below is a search section with 'Directories:' set to 'All Services' and 'Search Results:' showing 'SERVICE [1] Jan 22, 2009 12:59:57 PM'. A table below lists the results:</p> <table border="1"> <thead> <tr> <th>Name</th> <th>Center</th> <th>Type</th> <th>Status</th> <th></th> </tr> </thead> <tbody> <tr> <td>CaArraySvc</td> <td>CUMC</td> <td>DATA</td> <td>ACTIVE</td> <td>Query...</td> </tr> </tbody> </table>	Name	Center	Type	Status		CaArraySvc	CUMC	DATA	ACTIVE	Query...
Name	Center	Type	Status									
CaArraySvc	CUMC	DATA	ACTIVE	Query...								

Have a comment?

Please leave your comment in the [caArray End User Forum](#).