

---

# GridAnt: A Grid Workflow System

---

Kaizar Amin<sup>1,2</sup> and Gregor von Laszewski<sup>1</sup>

<sup>1</sup> Argonne National Laboratory, Argonne, IL 60439, U.S.A.

<sup>2</sup> University of North Texas, Denton, TX 76203, U.S.A.

## Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>GridAnt Tasks</b>	<b>2</b>
2.1	cogSetup . . . . .	2
2.2	gridAuthenticate . . . . .	3
2.3	gridExecute . . . . .	3
2.4	gridCopy . . . . .	5
<b>3</b>	<b>Installation</b>	<b>6</b>
<b>4</b>	<b>Examples</b>	<b>6</b>
4.1	gridExecute . . . . .	6
4.2	gridCopy . . . . .	7
<b>5</b>	<b>Workflow Example</b>	<b>7</b>

## 1 Introduction

The popularity of UNIX and DOS was initiated through the availability of sophisticated shells to express command line sequences. Soon it was noticed that complex workflow dependencies as part of the program generation process needed to be expressed. Makefiles have an essential part in automating the complex build process leading to a revolution in software engineering. It allowed users to structure their code and to build sophisticated reusable libraries with ease.

For the Java Platform a similar tool exists under the Apache Project with the name *Ant* [1]. It provides the Java community with a sophisticated build tool. Dependencies and parallelism can be expressed easily and new tasks can be included with little effort. As Ant is used by a large community it has been transitioned to a commodity tool and is the defacto standard for

controlling the build processes in Java. Every major Java interface development environment [4, 6, 8] includes the possibility to generate class files through Ant. Based on our experience with Grid workflows [9, 11] we propose to reuse the *Ant* framework to develop a simple yet powerful client side workflow system for Grids, which is named as *GridAnt* [10]. We find that the availability of the GridAnt framework provides a much needed functionality for testing and developing Grid applications with the different versions of the Globus Toolkit (GT2 and GT3) [5].

GridAnt proves to be an excellent tool, not only to map complex client-side workflows, but also as a simplistic client to test the functionality of different Grid services. GridAnt will help applications to make a smooth transition from GT2 to GT3. GridAnt is not claimed as a substitution for more sophisticated and powerful workflow engines [3, 2, 12]. Nevertheless, applications with simple process flows tightly integrated to work with the Globus Toolkit can vastly benefit from GridAnt without having to endure any complex workflow architectures. The philosophy adopted by the GridAnt project is to use the workflow engine available with Apache Ant and develop a Grid workflow vocabulary on top of it.

## 2 GridAnt Tasks

The following is a partial list of GridAnt tasks that we plan to implement.

- cogSetup : Grid environment setup via the Java Cog Kit.
- gridAuthenticate : Initializes the GSI-compliant proxy certificate to be used by clients.
- gridExecute : Executes an arbitrary job on a remote Grid resource.
- gridCopy : Provides third party file transfers between gridftp enabled Grid resources.
- gridQuery : Provides capabilities to query the Grid information servers.

This is a tentative list and is by no means final. Neither have we implemented the tasks for all the above mentioned services. The initial prototype for GridAnt has the functionality for cog setup, grid authentication, job submission, and file transfer. Other tasks are under development. We release the current version as a technology preview in order to obtain feedback and to engage the community in its further development.

### 2.1 cogSetup

The cogSetup task allows the user to set his Grid environment via the Java Cog Kit visual setup component (See Figure 1). It lets the user configure the Java CoG Kit by specifying various security related parameters, such as the locations of the user's long-term and proxy credentials, locations of the files containing trusted CA certificates, and some other options. The tool then creates a configuration file named *cog.properties*, which is used by the Java CoG Kit and GT3 security services. This task does not require any parameters.

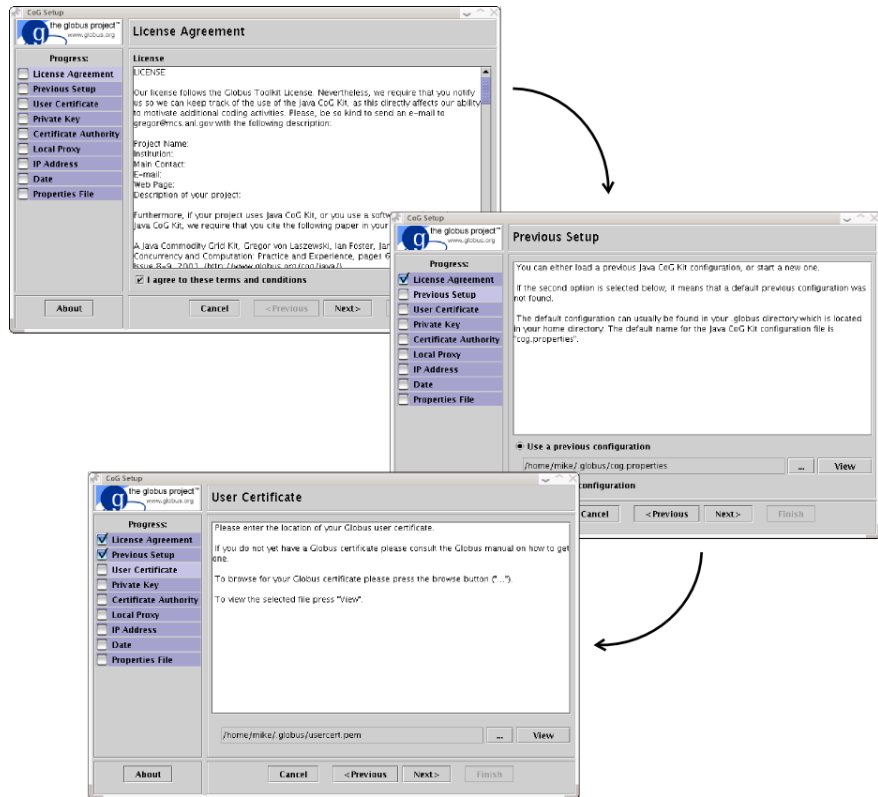


Figure 1: Grid Setup Wizard

## 2.2 gridAuthenticate

The gridAuthenticate task allows the user to visually generate his proxy certificate (See Figure 2). It allows the user to specify the proxy lifetime, cryptographic strength of the proxy, the locations of user's long-term credentials, and the location of the resulting proxy file. This task does not require any parameters.

## 2.3 gridExecute

The gridExecute task executes an arbitrary job on a Grid resource. It requires the following input parameters (\* specifies a mandatory argument).

- provider\* : Specifies the hosting environment for the Job Manager. Valid options are *GT2* and *GT3* for Globus Toolkit versions 2 and 3 respectively. The GridAnt client engages in a protocol based on the provider.
- server\* : Specifies the location of the Job Manager service available in the Globus Toolkit.
- security : Specifies the XML security parameters. Required only if the provider is *GT3*. Valid options are *xmlSig* and *xmlEnc* for XML signature and XML encryption respectively. The default is XML signature.
- delegation : Specifies the parameters for credential delegation for GSI security.

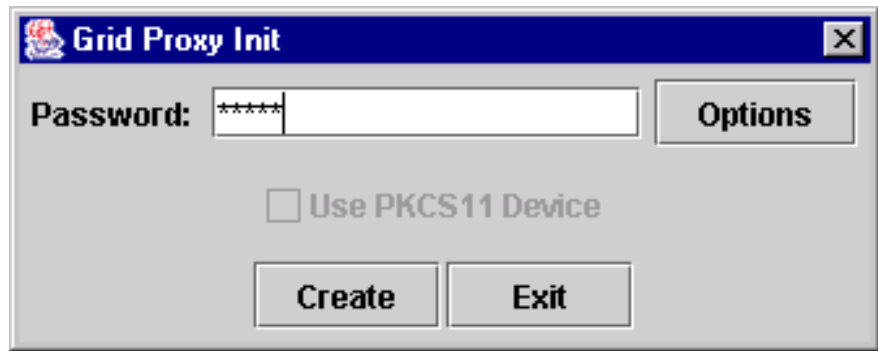


Figure 2: Proxy Initialization

Valid options are *full* and *limited* for full delegation and limited delegation respectively. The default is limited delegation.

- executable\* : Specifies the command to be executed on the Grid resource.
- localExecutable : A boolean flag that specifies if the executable resides on the client machine. The default is *false*.
- arguments : Specifies the arguments to be provided with the executed command
- directory : Specifies the directory in which the command is to be executed
- environment : Specifies the environment variables to be set prior to the execution of the command.
- outputFile : Specifies the file name to which the output must be redirected. If left blank, the output is streamed to the standard output. By default output is streamed to the standard output.
- errorFile : Specifies the file name to which the error must be redirected. If left blank, the errors are streamed to the standard error. By default the errors are streamed to the standard error.
- redirect : A boolean flag that specifies if the output and error streams are to be redirected to the client. Default value is *true*.

For example, if we like to schedule a job using provider GT3:

```

1 <gridExecute
2   provider="GT3"
3   server="http://hot.anl.gov:8080/.../SecureJobManagerFactory"
4   security="xmlEnc"
5   delegation="full"
6   executable="/bin/lis"
7   localExecutable="true"
8   arguments="-l"
9   directory="/home/user"
10  outputFile="myOutput.txt"
11  errorFile="myError.txt"
12  redirect="false"
13 />

```

or using provider GT2:

```

1 <gridExecute
2   provider="GT2"
3   server="hot.mcs.anl.gov:1234"
4   security=""
5   delegation="full"
6   executable="/bin/ls"
7   localExecutable="false"
8   arguments="-l"
9   directory="/home/user"
10  outputFile="myOutput.txt"
11  errorFile="myError.txt"
12  redirect="false"
13 />

```

## 2.4 gridCopy

The *gridCopy* task performs third party file transfers between grid resources capable of supporting the gridftp protocol. This task requires the following input arguments (\* specifies a mandatory arguments).

- provider\* : Specifies the hosting environment for the File Transfer. Valid options are *GT2* and *GT3* for Globus Toolkit versions 2 and 3 respectively. The GridAnt client engages in a protocol based on the provider.
- server\* : Specifies the location of the reliable file transfer factory service. Required only if the provider is *GT3*.
- security : Specifies the XML security parameters. Required only if the provider is *GT3*. Valid options are *xmlSig* and *xmlEnc* for XML signature and XML encryption respectively. The default is XML signature.
- delegation : Specifies the parameters for credential delegation for GSI security. Valid options are *full* and *limited* for full delegation and limited delegation respectively. The default id limited delegation.
- from\* : Specifies the url of the file to be copied. The url must be in the form `gsftp://machineName:portName/absolutePathName`.
- to\* : Specifies the url of the destination address. The url must be in the form `gsftp://machineName:portName/absolutePathName`.
- parallelStreams : Indicates the number of parallel tcp streams desired for the file transfer. Required only if the provider is *GT3*. Default is 1.
- tcpBuffer : Indicates the tcp buffer size desired for the file transfer. Required only if the provider is *GT3*. Default is 16384.

For example, if we like to transfer a file from machine *hot.anl.gov* to machine *cold.anl.gov* using provider *GT3*:

```

1 <gridCopy
2   provider="GT3"
3   server="http://ft.anl.gov:8080/.../ReliableTransferFactoryService"
4   security="xmlSig"
5   delegation="full"
6   from="gsftp://hot.anl.gov/home/user/from.txt"
7   to="gsftp://cold.anl.gov/home/user/to.txt"

```

```
8 | parallelStreams="3"
9 | />
```

or using provider GT2:

```
1 | <gridCopy
2 |   provider="GT2"
3 |   delegation="full"
4 |   from="gsiftp://hot.anl.gov/home/user/from.txt"
5 |   to="gsiftp://cold.anl.gov/home/user/to.txt"
6 | />
```

## 3 Installation

The current GridAnt framework has the following pre-requisites:

- Java 1.3.1. or higher. For the GridAnt framework to work with Java 1.4 it requires certain additional configuration for the new security libraries. If you intend to use Java 1.4.0 or higher , you will have to copy the Xalan.jar available in the gridant/lib directory to j2sdk1.4.0/jre/lib/endorsed/ directory.
- Jakarta Ant 1.5 or higher [1].
- GT3 alpha3 Server side components [7]. It is recommended that you use the comprehensive GPT binary installation available at

GT3 Alpha3 : <http://www.globus.org/ogsa/releases/alpha/index.html>

To install GridAnt you need to checkout the latest source code in the cvs repository.

```
1 | > mkdir cog
2 | > cd cog
3 | > cvs -d :pserver:anonymous@cvs.globus.org:/home/dsl/cog/CVS co -r
4 |   tp-1 gridant
5 | > cd gridant
6 | > ant build
```

## 4 Examples

Several examples are available in the build.xml file.

### 4.1 gridExecute

To test the *gridExecute* GridAnt task:

```
1 | > cd gridant
2 | > ant build
3 | > ... create your proxy certificate
4 | > ... start the GT3 service container (if provider is GT3)
5 | > ... edit the build.xml in the gridant directory such that the
6 |   arguments in that target "emph-gridExec" reflect the
7 |   appropriate values.
8 | > ant gridExec
```

## 4.2 gridCopy

To test the *gridCopy* GridAnt task:

```
1 > cd gridant
2 > ant build
3 > ... create your proxy certificate
4 > ... start the GT3 service container (if provider is GT3)
5 > ... edit the build.xml in the gridant directory such that the
6 arguments in that target "emph-gridCopy" reflect the appropriate
7 values.
8 > ant gridCopy
```

## 5 Workflow Example

To test the sample GridAnt workflow:

```
1 > cd gridant
2 > ant build
3 > ... create your proxy certificate
4 > ... start the GT3 service containers (if provider is GT3)
5 > ... edit the build.xml in the gridant directory such that the
6 arguments in that target "emph-workflow" reflect the appropriate
7 values.
8 > ant workflow
```

## Acknowledgment

This work was supported by the Mathematical, Information, and Computational Science Division subprogram of the Office of Advanced Scientific Computing Research, U.S. Department of Energy, under Contract W-31-109-Eng-38. DARPA, DOE, and NSF support Globus Project research and development.

## References

- [1] Ant – a Java-based Build Tool. <http://ant.apache.org>.
- [2] BPEL4WS: Business Process Execution Language for Web Services Version 1.0. <http://www-106.ibm.com/developerworks/webservices/library/ws-bpel>.
- [3] Condor DAGman. <http://www.cs.wisc.edu/condor/dagman>.
- [4] Eclipse: An Open and Extensible IDE. <http://www.eclipse.org>, 2003.
- [5] The globus project. <http://www.globus.org>.
- [6] JBuilder. <http://www.borland.com/jbuilder/>, 2003.
- [7] Open Grid Services Architecture (OGSA). <http://www.globus.org/ogsa>.
- [8] Sun ONE Studio. <http://www.sun.com/software/sundev/>, 2003.

- [9] Gregor von Laszewski. *A Parallel Data Assimilation System and Its Implications on a Metacomputing Environment*. PhD thesis, Syracuse University, December 1996.
- [10] Gregor von Laszewski, Kaizar Amin, Beulah Alunkal, Shawn Hampton, and Sandeep Nijure. Gridant – white paper. Technical report, Argonne National Laboratory, 2003. <http://www.globus.org/cog/grant.pdf>.
- [11] Gregor von Laszewski, Ian Foster, Jarek Gawor, and Peter Lane. A Java Commodity Grid Kit. *Concurrency and Computation: Practice and Experience*, 13(8-9):643–662, 2001.
- [12] Web Services Flow Language. [www.ibm.com/software/solutions/web-services/pdf/WSFL.pdf](http://www.ibm.com/software/solutions/web-services/pdf/WSFL.pdf).