

# How to deploy a parameter study in AstroGrid-D

G. Stöckle<sup>1</sup>

<sup>1</sup>Astronomisches Rechen-Institut (ARI)  
University of Heidelberg

AG Meeting Bonn 2010

# Outline

- 1 First steps in Grid Computing
  - Joining the Virtual Organisation (VO) AstroGrid-D
  - Setup and connection tests
- 2 The physical problem
  - Short description
  - Running the Code on the Grid

# Outline

- 1 First steps in Grid Computing
  - Joining the Virtual Organisation (VO) AstroGrid-D
  - Setup and connection tests
- 2 The physical problem
  - Short description
  - Running the Code on the Grid

# The Virtual Organisation AstroGrid-D

## What is AstroGrid-D?

- A project to set up Grid infrastructure for the German astronomical community,
- a collaboration of 14 institutes,
- part of the German D-Grid Initiative, using its infrastructure and standards,
- and using the Globus Toolkit (GT4) middleware .

## Some numbers:

- more than 100 users
- several compute nodes of 1000 cores total
- storage resources of over 100 TB

# How to join, step by step

1. Get an identification certificate from a higher instance (D-Grid ):

- Register\* at Grid Computing Centre Karlsruhe  
`https://gridka-ca-sec.fzk.de/` , the central certification authority for the german Grid
- your admin gets e-mail: do you approve this person is known to you?
- You get mail: *Ihr Zertifikat steht bei GridKa-CA zur Abholung bereit: [link](#)*
- fetch the certificate, copy\* to  
`$HOME/.globus/usercert.pem` and  
`$HOME/.globus/userkey.pem`
- import the certificate into your browser (!)

\*see in detail the Globus User Guide (on Slide "recommended reading").

# Membership registration in AstroGrid-D

## 2. Get access to AstroGrid-D:

- Go to `https://vomrs.gac-grid.org:8888` (using the Browser in which you just imported your certificate).
- Fill in the requested fields with your personal data → Receive an automatic E-mail-Confirmation of your data submission.
- An administrative of AstroGrid-D receives a mail with your request for membership and confirms it.
- You get a personal confirmation by mail.

NOW You are an approved member of AstroGrid-D!

# Outline

- 1 **First steps in Grid Computing**
  - Joining the Virtual Organisation (VO) AstroGrid-D
  - **Setup and connection tests**
- 2 **The physical problem**
  - Short description
  - Running the Code on the Grid

# Connecting to the Grid

Globus Toolkit is the middleware operating the Grid structure.

- Install GT4 from [1] (Even available for Windows).
- Call `Grid-Proxy-Init`: creates a proxy certificate used to identify yourself in AstroGrid-D.
- Login to a Grid-host using `gssissh GRIDHOST` (`gssissh`: Grid-Version of SSH). You need no password because your certificate is used by `gssissh`.

AstroGrid-D is a dynamical, changing and heterogenous network.

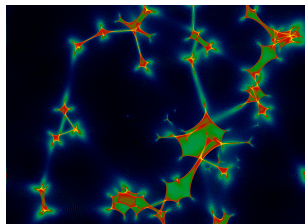
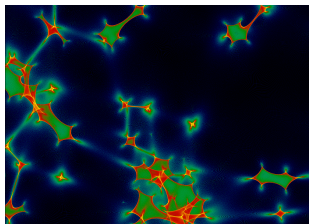
- Check `www.gac-grid.org` for an updated list of available resources.
- Try AstroGridTest [2] to test the resources, e.g. run `sitetest.sh` → connection status check or `submittest.sh` → testing to submit a job

# Outline

- 1 First steps in Grid Computing
  - Joining the Virtual Organisation (VO) AstroGrid-D
  - Setup and connection tests
- 2 The physical problem
  - Short description
  - Running the Code on the Grid

# A microlensing magnification map

Magnification caused by randomly distributed gravitational lenses: ray tracing + tree code (for gravitational potential approximation).



Changing parameter values lead to different maps + high resolution needed → computational times of many hours. Idea: Distributing the computation on multiple resources.

# A parameter grid

As parameters for the simulation can be chosen:

- surface mass density in stellar objects
- surface mass density in continuously distributed matter
- external shear

High number of rays in the ray tracing computation leads to very long computation times.

To distribute these computations we tried to distribute the computations on the resources of AstroGrid-D.

# Outline

- 1 First steps in Grid Computing
  - Joining the Virtual Organisation (VO) AstroGrid-D
  - Setup and connection tests
- 2 The physical problem
  - Short description
  - Running the Code on the Grid

# Code distribution

To compile Code and libraries on each machine (login and run make) is the easiest approach. Using svn and job description files, its also feasible to copy and compile code automatically on a certain ressource.

- Copy code using Globus-url-copy from local machine to grid machine (in this case mintaka):

```
globus-url-copy file://home/path/  
gsiftp://mintaka/gridpath/
```

- Compile code and make executable: `make ./microlens`

## creating different parameter files

A simple shell loop was written to create different parameter files containing different parameter values.

Serial computation approach:

- Copy Parameter files to the resource,
- Job start with `globusrun-ws -submit -s -c /bin/hostname`
  - `-submit`: job is submitted to a host directly or using an XML-based job description document
  - `-b`: Batch mode, enables to submit multiple jobs.
  - `-o`: creates an `.epr` file holding the current status of a job.
- Status check via `globusrun-ws-status`.
- When the job has finished, a new job can be started.

# Job description files

Start a job using an xml based .rsl job description file:

```
globusrun-ws -submit ... -f job.rsl
```

job.rsl:

```
<?xml version="1.0" encoding="UTF-8"?> <job>  
<executable>/MicrolensRun/microlens</executable>  
<stdout>/MicrolensRun/microlens.out</stdout>  
... </job>
```

This job description file allows schedulers to compile code on different resources, to run the code in a queued mode, also different monitoring portals need job description files.

# How to implement scheduling

- Job status can be checked and jobs can be distributed to faster machines or machines with less workload using `globusrun-ws -status`.
- To fully use the possibilities of the grid `globusrun-ws` jobs can be sent to a scheduler like Fork or Gridway . Gridway automatically handles data from different jobs and distributes them on different clusters inside a Virtual Organisation.
- `globus-run-ws ... -Ft gw`, turns on gw (Gridway) as scheduler.

# Summary

- Jobs can be submitted, files transferred and output collected using resources after registration in AstroGrid-D.
- Infrastructure is running and available. Scientific usecases have shown the usability of Grid Computing for Science (see [3]) .
- By reason of the abrupt ending of funding some software could not be finalised and exhaustively tested, anyway most developed tools can be of great value after understanding the involved concepts and problems.
- Code should be well written and understood, because automatic compilation of code in a Grid environment, perhaps even with the use of a Scheduler is much more sensitive to minor bugs then compilation on single nodes or on clusters.

# Outlook

- Grid Computing is more than just programming, its also a new way of communication and doing research.
- Resources can be shared independently of their geographical location and are no longer restricted to institute or university.
- Solutions can be identified, solved and used over the grid.
- Virtual Research Environments can be established.

# AstroGrid-D - recommended reading

- First steps in AstroGrid-D, Thomas Brüsemeister
- Running MPI Jobs on Grid Resources, AstroGrid-D Deliverable 1.4, H.Enke and Steve White, 06/2008
- app2grid 1.0 - Guide for porting User Applications to the Grid
- AstroGrid-D Test  
<http://www.rzuser.uni-heidelberg.de/~tg5/>
- Globus User Guide  
<http://www.gac-grid.org/project-products/grid-support/globus-user-guide.html>

# References



## Download Globus Toolkit

<http://www.globus.org/toolkit/downloads/>



**AstroGridTest SVN-Folder:** `svn://svn.gac-grid.org/software/AstroGridTest/trunk` **user: anonymous**  
**pwd: (empty)**



**AstroGrid-D: Grid Technology for Astronomical Science,**  
**H.Enke et al, New Astronomy, 07/2010**



<http://www.wissgrid.de>



<http://www.gac-grid.de>