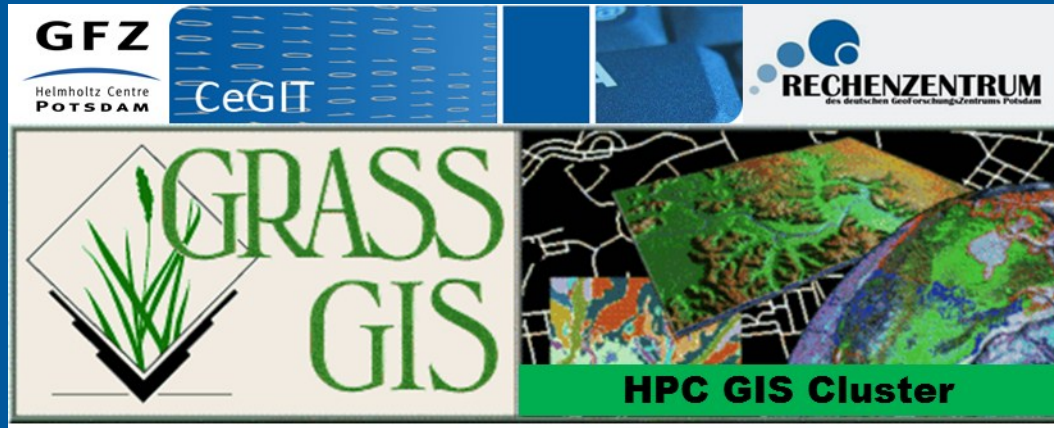


State of GIS at the High Performance Computing Cluster



Peter Löwe[°], Jan Thaler[°], Stefan Lüdtkke²

[°]Centre for GeoInformation Technology (CeGIT)

² Section 5.4 (Hydrology)

Computing Clusters

- A set of loosely connected computers that work together
- Which can be viewed as a single system
- Benefits over single computers of similar speed:
 - performance increase
 - no availability constraints
 - more cost- and energy-effective

Cluster GIS Workbenches

Benefits for geocomputation tasks

- **Parallelization** "Sorcerers apprentice"
- **Deployment of tasks with long duration** "Fire and forget"
- **Resource intensive tasks** "Size matters"
- **Secure and stable environment** "Murphies law"

The High Performance Computing Cluster

- 234 nodes / **480 CPUs** (AMD Opteron):
3084 cores
- **5 Tbyte RAM**
- **Load Sharing Facility (LSF) Batch-System** with multiple processing queues
- Suse Linux Enterprise 2011



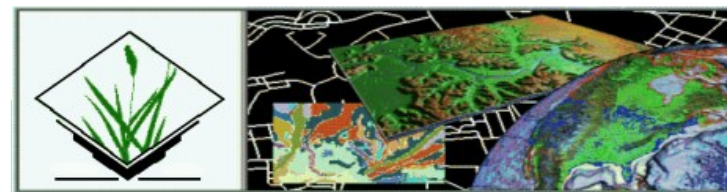
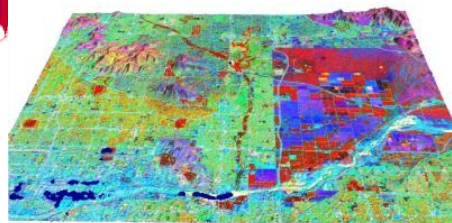
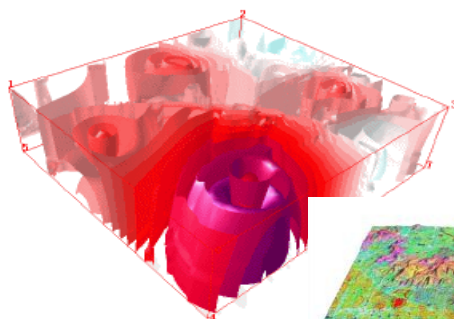
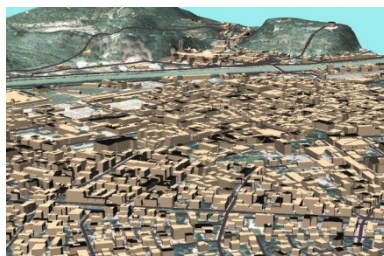
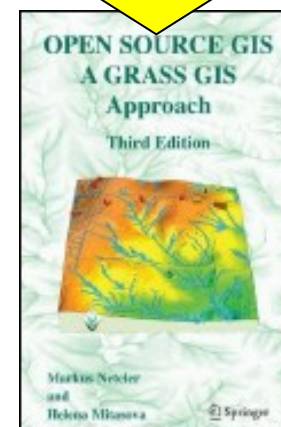
GIS (-related) Software on the HPC Cluster

- low level raster / vector processing: **Gdal/ogr**
- Desktop GIS:
 - **GRASS 6.4**
 - testing: **GRASS 7.0**
 - Highly experimental: **QGIS**
- Database Backends: **MySQL /SQLITE Databases**
- 3D Visualisation: **Paraview / NVIZ**
- Modelling / Statistics: **R**

Geographic Resources Analysis Support System (GRASS) GIS

- Largest and oldest Free and Open Source Software GIS
- > **350** modules + user contributed extensions
- Well documented, large user community
- Scripting: Shellscript, Python, Perl, etc.
- Compatible with various cluster environments

Available in
the GFZ
Library !



Images: <http://grass.osgeo.org/>

Knowledge Base: GFZ-Dokuwiki

<http://dokuwiki.gfz-potsdam.de/gis>

GIS-Cluster: GRASS GIS Dokumentation

This collection of links, documents and data provides information material for GRASS GIS users ranging from newbie to ninja and guru-level. Please be aware that the information is partially outdated, incomplete or recursive.

General Information

- [GRASS GIS Main Website](#)
- [GRASS 6.4 QUICK START](#)
- [GRASS 6.4 Graphical User Interface \(GUI\) Documentation](#)
- [GRASS GIS Mailing Lists](#)
- [GRASS GIS 6.4 development visualization from 1999 to 2011 with Gource](#)

GRASS Documentation

- [GRASS 6.4 Man Pages](#)
- [NVIZ 3D Visualisation Tool](#)
- [GRASS GIS WIKI \(english\)](#)
- [GRASS GIS WIKI \(german\)](#)
- [Books](#)
- [Tutorials and Courses](#)
- [Special Topics](#)

GRASS Layer Types

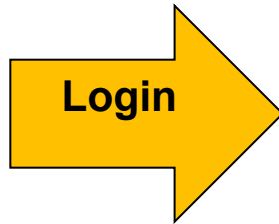
- [Introduction to Raster Data](#)
- [Introduction to Vector Data](#)
- [Introduction to Voxels / rast3d](#)

GFZ HPC Cluster Specific Information

- [First steps in 2008 \(german\)](#)
- [Overview Presentation EGU 2011](#)

How does this work ?

Sample Cluster GIS Session: Using PuTTY in Windows 7

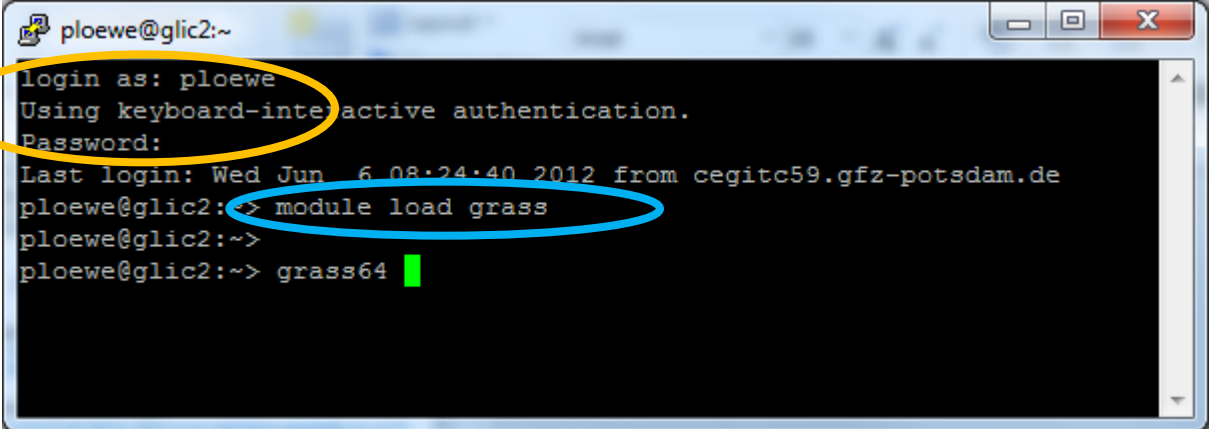


```
ploewe@glic2:~  
login as: ploewe  
Using keyboard-interactive authentication.  
Password:  
Last login: Wed Jun  6 08:24:40 2012 from cegitc59.gfz-potsdam.de  
ploewe@glic2:~> module load grass  
ploewe@glic2:~>  
ploewe@glic2:~> grass64 █
```

Sample Cluster GIS Session: Using PuTTY in Windows 7

Login

Load GRASS
module



```
ploewe@glic2:~  
login as: ploewe  
Using keyboard-interactive authentication.  
Password:  
Last login: Wed Jun  6 08:24:40 2012 from cegitc59.gfz-potsdam.de  
ploewe@glic2:~> module load grass  
ploewe@glic2:~>  
ploewe@glic2:~> grass64
```

Sample Cluster GIS Session: Using PuTTY in Windows 7

Login

Load GRASS
module

Start GRASS

```
ploewe@glic2:~  
login as: ploewe  
Using keyboard-interactive authentication.  
Password:  
Last login: Wed Jun  6 08:24:40 2012 from cegitc59.gfz-potsdam.de  
ploewe@glic2:~> module load grass  
ploewe@glic2:~>  
ploewe@glic2:~> grass64
```

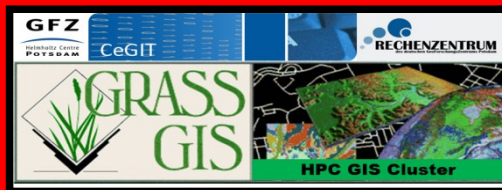
Sample Cluster GIS Session: Using PuTTY in Windows 7

Login

Load GRASS
module

Start GRASS

```
ploewe@glic2:~  
login as: ploewe  
Using keyboard-interactive authentication.  
Password:  
Last login: Wed Jun  6 08:24:40 2012 from cegitc59.gfz-potsdam.de  
ploewe@glic2:~> module load grass  
ploewe@glic2:~>  
ploewe@glic2:~> grass64
```



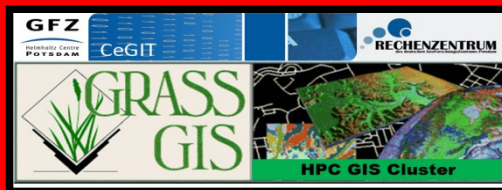
Sample Cluster GIS Session: Using PuTTY in Windows 7

Login

Load GRASS
module

Start GRASS

```
ploewe@glic2:~  
login as: ploewe  
Using keyboard-interactive authentication.  
Password:  
Last login: Wed Jun  6 08:24:40 2012 from cegitc59.gfz-potsdam.de  
ploewe@glic2:~> module load grass  
ploewe@glic2:~>  
ploewe@glic2:~> grass64
```



```
ploewe@glic1:~  
Main Options VT Options VT Fonts  
  
GRASS GIS  
  
Welcome to GRASS 6.5.svn (2011)  
GRASS homepage: http://grass.osgeo.org/  
This version running thru: Bash Shell (/bin/bash)  
Help is available with the command: g.manual -i  
See the licence terms with: g.version -c  
If required, restart the GUI with: g.gui tcltk  
When ready to quit enter: exit  
GRASS 6.5.svn (rapideye_elhierro):" > █
```

Command Line

Sample Cluster GIS Session: Using PuTTY in Windows 7

Login

Load GRASS
module

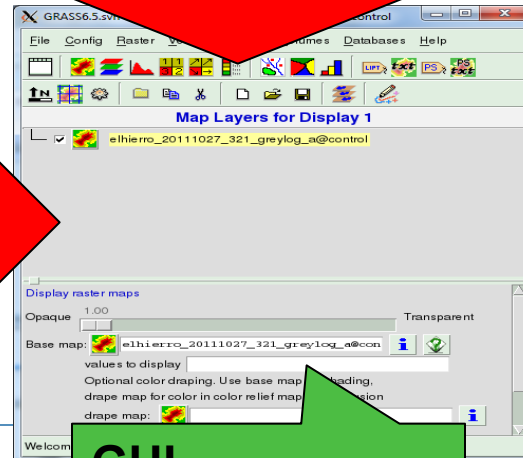
Start GRASS

```
ploewe@glic2:~  
login as: ploewe  
Using keyboard-interactive authentication.  
Password:  
Last login: Wed Jun  6 08:24:40 2012 from cegitc59.gfz-potsdam.de  
ploewe@glic2:~> module load grass  
ploewe@glic2:~>  
ploewe@glic2:~> grass64
```



```
GRASS GIS  
Welcome to GRASS 6.5.svn (2011)  
GRASS homepage: http://grass.osgeo.org/  
This version running thru: Bash Shell (/bin/bash)  
Help is available with the command: g.manual -i  
See the licence terms with: g.version -c  
If required, restart the GUI with: g.gui tcltk  
When ready to quit enter:  
GRASS 6.5.svn (rapideye_elhiero):" > |
```

Command Line



GUI

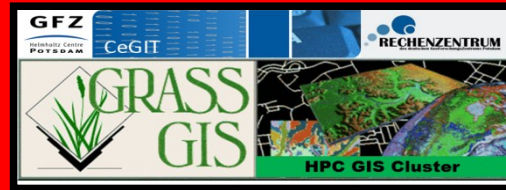
Sample Cluster GIS Session: Using PuTTY in Windows 7

Login

Load GRASS
module

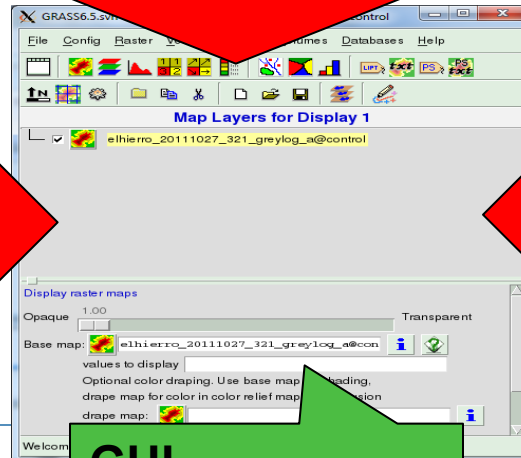
Start GRASS

```
plowe@glic2:~  
login as: plowe  
Using keyboard-interactive authentication.  
Password:  
Last login: Wed Jun  6 08:24:40 2012 from cegitc59.gfz-potsdam.de  
plowe@glic2:~> module load grass  
plowe@glic2:~>  
plowe@glic2:~> grass64
```

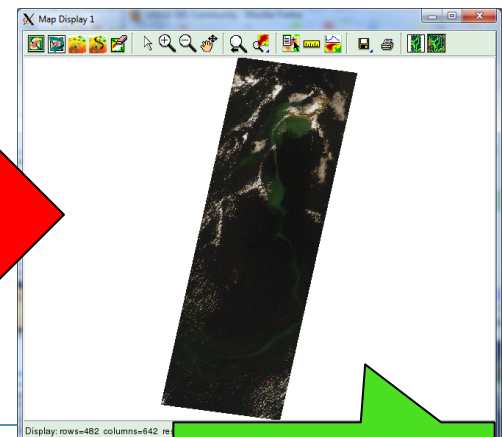


```
plowe@glic1:~  
Main Options VT Options VT Fonts  
  
GRASS GIS  
  
Welcome to GRASS 6.5.svn (2011)  
GRASS homepage: http://grass.osgeo.org/  
This version running thru: Bash Shell (/bin/bash)  
Help is available with the command: g.manual -i  
See the licence terms with: g.version -c  
If required, restart the GUI with: g.gui tcltk  
When ready to quit enter: exit  
GRASS 6.5.svn (rapideye_elhiero):" > |
```

Command Line



GUI



Map
Window

Latest improvements

- GRASS6.4.2 (stable)
 - NetCDF Support
 - Databases: MySQL
- GRASS7 (devel)
 - New enhanced wxGui
 - **Temporal data sets** (and queries) over vector, raster and volume data (or combinations)
- GRASS / Cluster Job dispersal system
 - Job invocation right from the commandline (new)
 - Via GRASS-Scripts (traditional)

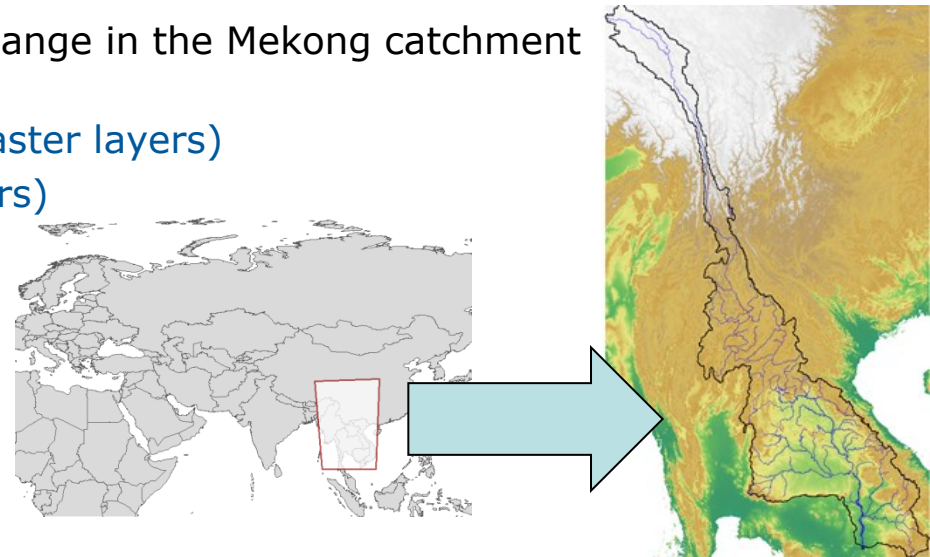
New: The GIS Servicedesk

gis-servicedesk@gfz-potsdam.de

- New unified point of contact for GIS-related support, provided by CeGIT:
 - Handles ESRI GIS / FOSS GIS
 - Software installation/ access
 - Bug Reports
 - Wish list
 - Request special features
 - „Ask a Guru“

GIS Usage in the Cluster

- Centre for GeoInformation Technology (CeGIT)
 - Large Scale Vector Processing (Tsunami Simulation)
 - Globe Maps / Rendering
 - 2D animations
- Department 5 – Hydrology
 - WISDOM Project:
 - Sediment dynamics and Climate Change in the Mekong catchment ($\sim 800.000 \text{ km}^2$)
 - Managment of climate data (~ 50.000 raster layers)
 - Precipitation data (~ 21.000 raster layers)
 - Statistics



Maps: S. Lütke, GFZ, 2012

Challenge: World Fair EXPO 2012



SEA EVOLUTION 바다의 친환경적 진화

German Pavilion Expo 2012 Yeosu

2012 여수세계박람회의 독일관

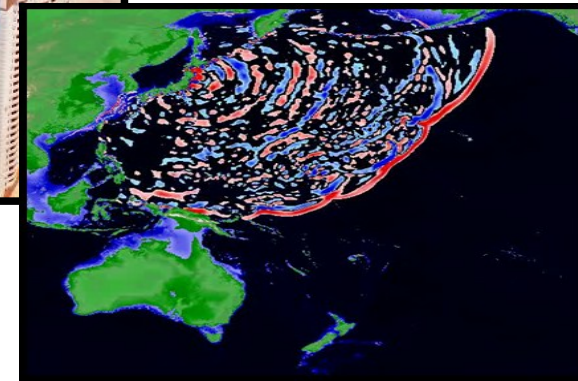
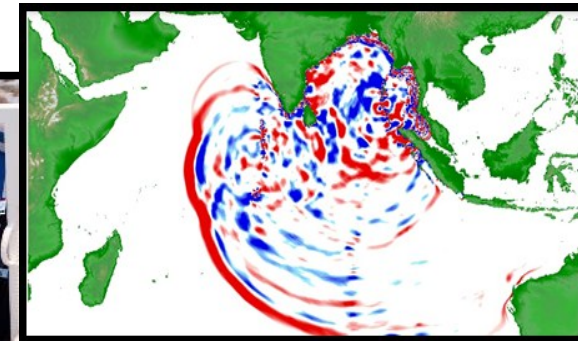
<http://expo2012-deutschland.de/de/>



<http://expo2012-deutschland.de/de/>



<http://www.gfz-potsdam.de/portal/cms/GFZ/Public%20Relations/Pressemitteilungen/2012/120615-GitewsExpo>



GFZ was invited to create tsunami animations for the german pavillion at the EXPO 2012 in Yeosu, Korea.

This task was completed using the GFZ High Performance Computing Cluster.

Stats: Tohoku-Tsunami Animation

- Parallel execution:
 - Derivation of **751** thematic maps based on Tsunami simulations
 - Executed in parallel cluster jobs
 - Job processing: **250 - 280 seconds** each (~ 5 min)
 - Overall processing time for **751** jobs: **68 Minutes.**
- Guestimate for serial execution (on similar hardware):
 - **751 Jobs** taking **250 seconds** each => **about 3219 minutes**
 - = **about 52 hours**
- Reality Check: All the Tsunami related work in 2012 resulted in **0,076 % of the annual cluster load**

Conclusion: A LSF-based HPC GIS Workbench

- A research utility in the sense of “Software as a Service” (SaaS)
- A first step towards building a GFZ corporate cloud service.
- Allows to tackle GIS tasks previously out of reach of conventional workstations.
- High performance geocomputation becomes available for an audience beyond conventional HPC / Grid power users.

All GFZ researchers are welcome to use this service offer

gis-servicedesk@gfz-potsdam.de