1. Introduction

In many fields of science and industry the need for computing resources is permanently growing. New technologies are developed to fulfill those needs and requirements. Grid computing is a promising way to cope with the upcoming computing challenges [1]. Scientists of various fields try to benefit from the concept of grid computing and as a result more and more grid communities are founded.

The Georg-August-University of Göttingen in Germany with its research campus of university and non-university research institutions hosts many established and new grid communities. The idea of sharing compute resources, personnel, and expertise appeared obvious and lead to the GoeGrid project, a joint resource center of the computer center GWDG [2], high energy physics community [3], MediGrid [4], TextGrid [5], and the department of physics of the University of Göttingen [6].

2. GoeGrid Computing Center

In the following the communities and organizations participating in the GoeGrid project and their research fields are briefly introduced:

2.1. GWDG

The GWDG (Gesellschaft für wissenschaftliche Datenverarbeitung mbH Göttingen) serves as a computer and competence center for the Max Planck society and the University of Göttingen.
The GWDG hosts the GoeGrid resource center, and it is one of the participating grid communities. One of its projects is called Optinum-Grid [7]. Its aim is to optimize technical systems and scientific models with the help of numerical simulations on the grid. The GWDG maintains the hardware, the network infrastructure, and the water-cooling of the racks.

2.2. Biomedicine
The MediGrid and its follow-up projects Services@MediGrid and WissGrid apply grid computing for biomedical research, image processing, and clinical research. They are part of the D-Grid project which is the German Grid Initiative [8]. With the D-Grid project the Federal Ministry of Education and Research (BMBF) is funding the development of several grid projects of German research and industry. One of the aims of MediGrid is the integration of the grid middleware embedded in a user-friendly environment that allows simple access to e-Science services for biomedicine. The three services are based on an enhanced security model with different security level. It is possible to perform computations on medical data and pseudonymized patient data with access to special secured resources [9, 10]. The project is coordinated from Göttingen.

2.3. TextGrid
Like MediGrid, TextGrid is part of the D-Grid project to establish a grid-based infrastructure for science and industry in Germany. TextGrid introduces e-Science concepts and technologies into the humanities, first providing a comprehensive toolset for researchers in the philologies, linguistics, and related fields. The TextGridLab establishes a virtual research environment for the collaborative editing, annotation, analysis and publication of specialist texts. The grid provides standard ways for archiving digital objects (digitized manuscripts, annotated XML full texts, etc) and as a basis for collaborative research by scholars from all over the world. It also initiates the creation of an open service environment for sharing tools and workflows. TextGrid initially consists of ten institutions, domain experts and grid technologists. Aiming to extend TextGrid into a sustainable e-Humanities infrastructure, the consortium invites new partners and establishes ties with international e-Humanities initiatives.

2.4. Department of Physics
The department of physics of the University of Göttingen uses the resources of GoeGrid for CPU intensive simulations. The main user is the institute of theoretical physics. Its solid state physics group uses state-of-the-art computer algorithms, like quantum Monte Carlo or numerical renormalization group schemes. Another group studies statistical physics of soft matter by computer simulations. The experimental physics group in high energy particle physics and their applications are described in the following.

2.5. High Energy Physics
High energy physics (HEP) aims at understanding the most fundamental constituents of matter and their interaction. The main experimental facility used by modern physics to answer such questions is the Large Hadron Collider (LHC), a proton-proton collider located at CERN in Geneva, Switzerland. It is designed for particle collisions at energies up to 14 TeV center-of-mass energy. The LHC collision data are investigated with huge detectors. There are four large experiments at LHC: two multi-purpose experiments called ATLAS and CMS, and the two specialized experiments ALICE and LHCb. At the design luminosity and bunch crossing rate of 400 MHz, the expected event rate is of the order of $10^9$ Hz. The ATLAS detector has about 100 million readout channels to detect collision events. This leads to a data rate of 1 TB/s.
This data rate has to be reduced due to limitations set by processing time and storage capacity. The total trigger rate reduction provided by the ATLAS trigger system is about $10^7$, which corresponds to 100 Hz of event rate for permanent storage, while the size of a single event is about 1.6 MB. This leads to a transfer rate of 300 MB/s. The four experiments are expected to produce 15 PB of data per year.

The high energy physics community makes extensive use of world wide available grid computing resources. The II. Physikalisches Institut of the University of Göttingen is member of the international ATLAS collaboration as well as the WLCG collaboration. Goegrid participates in the world wide processing of ATLAS data and provides a so-called Tier-2 (regional) and Tier-3 (local) center as described in section 4.

### 3. Cluster Resources and Architecture

#### 3.1. Hardware Setup

The Goegrid cluster provides 1748 CPU-cores. Additional 672 cores will be installed in December 2010. Table 1 lists the different types of hardware being used. There are 2 – 3 GB of RAM per core available.

| CPU               | Clock [GHz] | RAM [GB] | hard disk [GB] | quantity |
|-------------------|-------------|----------|----------------|----------|
| 2 Intel Xeon X5355 QuadCore | 2.66      | 16       | 1500           | 30       |
| 2 Intel Xeon X5355 QuadCore | 2.66      | 16       | 292            | 78       |
| 2 Intel Xeon 5160 DualCore | 3.0       | 8        | 400            | 13       |
| 2 × 2 Intel Xeon E5450 QuadCore | 2.83 | 2 × 16  | 2 × 240        | 48       |
| 2 × 2 Intel Xeon E5530 QuadCore | 2.4      | 2 × 24  | 2 × 300        | 8        |
| 2 Intel Xeon X5650 SixCore | 2.66      | 36       | 1000           | 56       |

Each Goegrid community has different needs and requirements for mass storage. Two different approaches are chosen for Goegrid called dCache [11] and Storage Area Network (SAN). The high energy physics community uses local storage servers with a total disk space of 516 TB. The storage is managed by dCache, which is a storage system adopted by many HEP grid sites. The dCache project provides a system for storing and retrieving huge amounts of data, distributed among a large number of heterogeneous server nodes, under a single virtual filesystem tree with a variety of standard access methods.

The other Goegrid communities use the SAN. The advantage of SAN is the virtualization of
the disk space, i.e. all disks are combined to one virtual storage area. The virtualization allows to move or mirror files transparently to the user. The SAN is maintained by the GWDG. In addition to the disk space the tape system of the GWDG provides additional 30 TB. Table 2 summarizes the mass storage of GoeGrid.

The requirements of local disk space for the execution of jobs differ among the GoeGrid communities. This results in a non-uniform node configuration in the cluster. There are 30 nodes configured with 1.5 TB of local disk space (see Table 1), which is required by some applications of the theory department. These nodes can be used by all communities, however the output intensive jobs of the theory users can only run on these dedicated nodes. Newer nodes have larger local storage, as well as the requirements for local disk space also rises for the ATLAS experiment.

### Table 2. Mass storage of the GoeGrid communities.

| Community              | Type   | Disk Space [TB] | Tape [TB] |
|------------------------|--------|-----------------|-----------|
| High energy Physics    | dCache | 516             | -         |
| TextGrid               | SAN    | 280             | 280       |
| Department of physics  | SAN    | 72              | -         |

3.2. Software Setup

For the interoperable use of the GoeGrid resources all communities agreed on a common setup. This setup meets requirements of the various applications. The major decisions made were the choices of the grid middleware and the operating system of the compute nodes. The high energy physics requirements are given by the WLCG, the grid computing organization for HEP described in the next section. The middleware being used is gLite 3.1 and 3.2 [12]. The GoeGrid project started as a bottom-up approach. All communities agreed on the installation of CentOS 5 linux. This operating system is compatible with all community specific criteria and applications and middlewares. This is crucial for the interoperable use of compute nodes. GoeGrid is accessible via gLite and the Globus Toolkit.

The installation and updates of all needed software packages on the compute nodes are managed by the cluster management tool Rocks [13]. This tool allows simple and fast changes of the configurations of all nodes in the cluster with the management tool cfengine [14]. Rocks also includes the monitoring tool ganglia [15, 16] (see section 6). The local resources are assigned to the jobs by the batch system torque with the scheduler maui.

4. Worldwide Large Hadron Collider Computing Grid

The Worldwide LHC Computing Grid (WLCG) is the global organization to maintain the data storage and analysis infrastructure for the four LHC experiments. More than 200 computing centers in 34 countries participate.

The WLCG computing model defines the strategy on how to process and distribute the experiment data. There is a hierarchical structure with centers organized in tiers. At CERN, where the collision data are produced, there is the so-called Tier-0 (global) center. It serves as mass storage system for all the raw data of the experiments. A first pass of the event reconstruction and calibration is done as well. From the Tier-0 center, the data are further shipped to several national Tier-1 centers. At Tier-1 centers, the data reconstruction can be rerun with improved calibration and reconstruction algorithms. In addition, permanent storage
for data is provided. Data from the Tier-1 are further replicated to regional Tier-2 centers. They continuously run simulation jobs for Monte Carlo production, and user-specific analysis jobs and provide storage for researchers and for physics groups. Finally, Tier-3 centers are used for data analysis and storage for local users.

5. GoeGrid as ATLAS-Tier-2 Center
In 2010 the global analysis of the initial LHC data started. The GoeGrid cluster is heavily involved in the ATLAS Monte Carlo production and offers its resources for user analysis of the LHC data. GoeGrid provides stable and sustainable operations during constant growth of the cluster usage. In Figure 1, the elapsed time (walltime) of all Monte Carlo production jobs and user analysis jobs per month is shown. The constant increase of activity is correlated to the LHC data taking. Another way to see the rising analysis activity in 2010 is to look at the user storage. In Figure 2, the reserved and used disk space for single users can be seen as a function of time. This disk space is part of the dCache storage at GoeGrid.

![Figure 1. Walltime of production and user jobs.](image)

![Figure 2. Total and used disk space for users at GoeGrid.](image)

6. Cluster Monitoring
For the stable operations of a Tier-2 center it is crucial to have an efficient multi-level monitoring system to obtain detailed diagnostics of possible malfunctions of the cluster. GoeGrid utilizes several tools to monitor jobs from grid-level down to the level of hardware components. Some of these tools are described in the following:

6.1. Hardware Monitoring
Hardware vendors typically provide hardware monitoring tools with detailed low-level information on the hardware state such as CPU temperature or memory failures amongst other things.

6.2. Ganglia
Ganglia is an open-source linux monitoring tool for clusters. It allows to display various metrics of nodes like CPU, disk, memory, and network usage. The information is clearly arranged on a web page. The displayed graphs are produced with the rrdtool [17].
6.3. Nagios

Nagios is a professional open-source monitoring tool [18] for host based monitoring tests. A set of services to be tested is defined for each host. The nagios server collects the results of all tests and displays them on a detailed web page. Tests are implemented as plugins which allow to easily extend nagios for special purposes. Lots of predefined plugins exist to test services of the operating system as well as grid services. Nagios can be configured to notify administrators about possible problems, e.g. via email or mobile phone. The logic when to trigger an alarm can be customized.

6.4. HappyFace

In addition to the site-specific monitoring web pages, there are lots of experiment specific tools. It can be very tedious to check all the sources of information relevant for a site. This problem lead to the so-called HappyFace project [19]. HappyFace is a meta-monitoring tool. It summarises monitoring information from arbitrary sources and displays a summary of the status on a web page. It is a plugin-based and easily extendable framework. The plugins are grouped in categories and the state of a single test or a category is displayed by a smiley. All the monitoring information is stored in a database allowing to check the history of certain monitoring tests. The project is a joint effort of the universities in Aachen, Göttingen, Hamburg, and Karlsruhe in Germany and is supported by the Helmholtz-Alliance “Physics at the Terascale”.

7. Conclusion

GoeGrid is an inter-disciplinary operated computing resource center. The participating communities are actively involved in the administration of the cluster. This allows to solve all the challenges originating from the different requirements of each community. The largest partner is the high energy physics group. Its resources are integrated in the world-wide LHC computing as an ATLAS Tier-2 center. There is a high activity of ATLAS user and production jobs and the usage is constantly growing. The effective management of the cluster using various monitoring tools provides stable cluster operation for the D-Grid and WLCG communities.

References

[1] Foster I and Kesselman C 1999 The Grid: Blueprint for a New Computing Infrastructure (Morgan Kaufmann publishers)
[2] GWDG, http://www.gwdg.de/
[3] II. Physikalisches Institut, University of Göttingen, http://physik2.uni-goettingen.de/
[4] MediGrid, http://www.mediagrid.de
[5] Küster M W, Ludwig C, Neuroth H and Aschenbrenner A 2007 IEEE DEST, Cairns, Australien, SPECIAL SESSION 3: e-Humanities for Digital Eco-systems: A Social, Cultural, Economic and Political Agenda.
[6] Department of Physics, University of Göttingen, http://physik.uni-goettingen.de/
[7] Optinum-Grid http://www.d-grid-gmbh.de/index.php?id=129
[8] D-Grid Initiative, http://www.d-grid.de
[9] Kotthä S, Peter K, Steinke T, Bart J, Falkner A, Viezens F, Mohammed Y, Sax U, Hoheisel A, Ernst T, Sommerfeld D, Krefting D and Vossberg M 2007 German e-Science Conference
[10] Sax U, Mohammed Y, Viezens F and Rienhoff O 2007 Grid-Computing in der biomedizinischen Forschung (Urban & Vogel)
[11] DCache, http://www.dcache.org
[12] GLite, http://glite.web.cern.ch/glite/
[13] Rocks cluster management, http://www.rocksclusters.org
[14] Cfengine http://www.cfengine.com/
[15] Sacerdoti F, Katz M, Massie M and Culler D E 2003 IEEE Cluster 2003 Conference, Hong Kong
[16] Massie M L, Chun B N and Culler D E 2004 Parallel Computing 30
[17] Rrdtool http://www.mrtg.org/rrdtool/
[18] Nagios, http://www.nagios.org/
[19] HappyFace Project, https://ekptrak.physik.uni-karlsruhe.de/trac/HappyFace/wiki/Version_2/