Data Grid Tutorials with Hands-on Experience*

Heinz Stockinger[1], Flavia Donno[1,2], Roberto Puccinelli[3], Kurt Stockinger[1]
{Heinz.Stockinger|Flavia.Donno|Kurt.Stockinger}@cern.ch
r.puccinelli@cedrc.cnr.it
[1]CERN, European Organization for Nuclear Research, Geneva, Switzerland
[2]INFN, National Institute for Nuclear Physics, Italy
[3]CNR, Italian National Research Council, Rome, Italy

Abstract

Grid technologies are more and more used in scientific as well as in industrial environments but often documentation and the correct usage are either not sufficient or not too well understood. Comprehensive training with hands-on experience helps people first to understand the technology and second to use it in a correct and efficient way. We have organised and run several training sessions in different locations all over the world and provide our experience. The major factors of success are a solid base of theoretical lectures and, more dominantly, a facility that allows for practical Grid exercises during and possibly after tutorial sessions.

1 Introduction

Since the creation of the Grid vision in 1998 [14], tremendous progress has been made in building and providing Grid technologies to researchers and application users. Many Grid projects have been started all over the world that provide all kinds of software solutions to academia and industry. In order to promote outstanding new technologies as well as their efficient usage, comprehensive training is required that covers several methodologies and usage aspects. Today, Grid technologies still require detailed knowledge of internal system structures and machine dependencies. Grid software installation is often rather difficult and requires several hardware resources, which makes it difficult for conventional users to quickly gain experience with Grid technology.

Although many end users have already heard about the basic Grid concepts, often the detailed usage is not clear to them. Basic architectural features that are the result of the distributed nature of the Grid have to be explained in detail in order to help end users build the correct mental model for understanding and using Grid software. Introductory or overview talks and lectures are provided by several of the Grid development projects but often hands-on training is missing.

Within the EU DataGrid project [7] (EDG), we started in 2002 to provide comprehensive training in EDG tutorials [26], where people learn about Grid computing (in particular Data Grids), the EU DataGrid project and all the software components provided to the end user. In addition, participants gain practical experience with the software systems provided by EDG during dedicated hands-on sessions where tutors assist the students with any questions or problems they might have. We believe that such hands-on training is more effective to end users and they profit from it more than from simple presentations or short demonstrations. In addition, hands-on sessions provide a realistic view on what can be achieved with Grid technologies today and where potential problems lie. Feedback is immediately forwarded to our developers and thus the tutorials are an integral part of the software development cycle.

One of the critical points of such a tutorial is a Grid infrastructure used by the students as a training facility. Originally, we used the EDG application testbed which includes several sites all over Europe. However, it turned out that tutorials load the system considerably such that project internal application users often had to stop their work in order to give preference to tutorial users. Therefore, we created a testbed dedicated to dissemination and training events such as tutorials. We report on the experiences of running such a tutorial and what is necessary in order to make them effective for end users.

We believe that we provide a unique tutorial programme in which more than 500 people world-wide have already participated. The tutorial is now simple to prepare (including practical exercises) and can easily be adopted to any

*This work was partially funded by the European Commission program IST-2000-25182 through the EU DataGrid Project as well as INFN through the INFNGrid project

kind of classroom teaching in universities or public events as our experience indicates.

The document is organised as follows. Section 2 provides an overview of the tutorial programme (theoretical and practical parts). Since the software system and its preparation are major parts of the tutorial, we give details in Section 3 of the installation, configuration and what is required to run hands-on exercises. More detailed general experience with tutorials is given in Section 4 and general teaching aspects are discussed in Section 4.3. Related work on tutorials is presented in Section 5.

2 Tutorial Programme

In the following section we give an overview of the basic tutorial programme we have been teaching for more than a year. The current version described below is aimed primarily at applications users, and it provides the fundamental theoretical background to understand the main software components and their interactions.

2.1 Lecture Programme

The tutorial programme consists of nine theoretical lectures which introduce the students to Grid computing in general, discuss the EDG software components and their deployment on the EDG testbed, show examples of how application groups successfully exploit Grid computing for solving their problems, and give an outlook on future directions of Grid computing. For each lecture we list the key concepts and we refer the reader to the tutorial web-site [6] (where all talks are listed) for further information.

1. **Introduction to Grid Computing & EU DataGrid Project**: An overview is given of Grid computing in general and the EDG project. The project organisation and the general software architecture are described, and some related projects are outlined.

   Key concepts: Grid Computing, Data Grids, international Grid projects world-wide, EU DataGrid project

2. **Security Issues**: Secure access to Grid resources is a major issue and security is one of the first things a user has to deal with when starting to use the Grid. A basic overview about current security solutions in the EDG project is given.

   Key concepts: GSI Security (*Grid Security Infrastructure* provided by the Globus Alliance), user and host certificates, Virtual Organisation Management

3. **Testbed Overview**: EDG deploys a large-scale testbed that spans several sites all over Europe. Definitions are given of the available Grid services and resources and a detailed overview of the testbed is provided.

   Key concepts: logical machine types (User Interface, Storage Element, Computing Element, Worker Node, Information Service, Resource Broker, etc.), overview of EDG's international testbed, interoperability

4. **Workload Management**: Most EDG users interact with the EDG software system by submitting their jobs (executable programs) to a Resource Broker which performs matchmaking on available and requested resources and then dispatches jobs to resources in the testbed. This lecture provides background information on the workload management software system and details about job submission.

   Key concepts: user interaction with Workload Management System (WMS), components (Resource Broker, Logging & Bookkeeping, etc.), Job Description Language (JDL)

5. **Data Management**: One of the main objectives of a Data Grid is the management of large distributed data stores. This lecture gives an overview about replica and metadata management as well as the software tools provided by EDG to deal with data management problems.

   Key concepts: replica management system, Replica Location Service (including Replica Metadata Catalogue), Replica Access Optimisation, Storage Resource Management

6. **Information Service**: In a Grid environment, there are several hardware and software resources that can be used by end-users as well as Grid services and applications. Information systems are used to keep track of resources and also to monitor the current status of the Grid. The EDG solution and the user interaction with it is outlined in detail.

   Key concepts: Relational Grid Monitoring Architecture (R-GMA), Consumer, Producer, Registry, Archiver, Glue Schema

7. **Software Installation/Configuration**: This lecture gives a brief introduction on how to obtain, install and configure the EDG software tools.

   Key concepts: LCFG (a software installation and configuration tool), EDG software repository, fabric management

8. **Applications**: In the EDG project, three major application domains are supported: High Energy Physics, Earth Observation and Biomedical Applications. This talk gives a brief overview of these applications and how they use Grid tools.

Key concepts: large scale data processing, batch versus interactive analysis

9. **Future Direction**: This lecture briefly covers the future of the EU DataGrid project as well as Grid computing in general.

   Key concepts: SOAP, OGSA, Grid Services, Web Services

A graphical overview of the lectures with their dependencies is given in Figure 1. Each of the lectures is between 30 and 45 minutes long and in a typical setup lectures 1-4 are given in the first day, and lectures 5-9 during the second day. Lectures are typically followed by hands-on exercises (see Section 2.2) which allow the students to get real experience with the topics covered in the lectures. The lecture material is presented in such a way that no specific background on Grid technology is required, and all basic details for the usage in the hands-on exercises as well as in potential real-world applications is given. A filmed version of the tutorial lectures is available and can be requested from the EDG dissemination group.

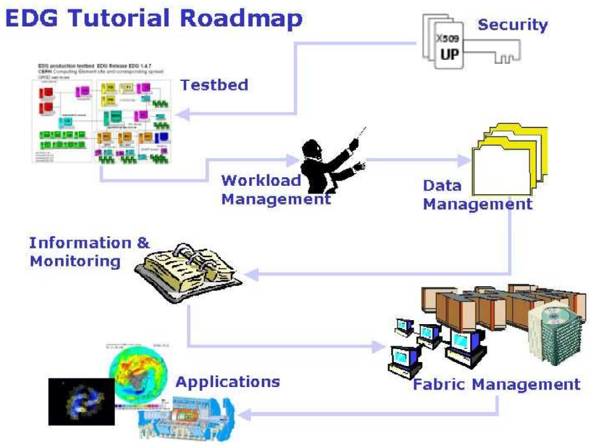

Figure 1. Overview of the lecture programme

2.2 Practical Hands-on Exercises

One of the main goals of the tutorial programme is to give students hands-on experience with EDG and the underlying Globus software on a distributed Grid testbed. Usually, the Grid Dissemination Testbed (GriDis) [16] is used for this purpose. GriDis hosts the main Grid infrastructure and can be extended to include further sites running the required software version (LCG-1). In order to increase the resources of the testbed, sites from other projects, in particular from the EU CrossGrid [5] project, can temporarily be added to the testbed available to the students.

The exercises are typically performed on client machines or laptop computers from which the students log in to the testbed's *User Interface*, the gateway to the Grid which hosts all the client software required to interact with the Grid. Using laptop computers has proven to be very successful since no special hardware nor operating system is necessary to use the Grid. In principle, all what is required to do the exercises are an ssh client and potentially a web browser.

In the hands-on session we focus on the following three aspects: job submission, data management and information systems. For each of these areas, students are given exercises [10] with the respective solutions on how to use the Grid with both command line tools and C++ or Java APIs. This allows the students to get experience with the real usage of a Grid environment. Students are free to go at their own pace in the afternoon sessions and tutors help in case of questions or problems.

The hands-on exercises are organised in a way that students learn about all the major points that are necessary to use Grid technology. The first obstacle is usually to get registered on the Grid using a user certificate. All user certificates are prepared and usually pre-installed on the client machine that is used to access the Grid. Thus, the students can see how to use the Grid security infrastructure and what potential problems exist.

Once the security exercise is done, students can start with job submission exercises which have different levels of difficulty. We have chosen basic and advanced exercises in order to address our diverse user community. The basic exercises contain all solutions in the exercise hand-out and people are requested to follow the written instructions. Advanced exercises usually require more profound knowledge about the work load management system. Since all user manuals are available on line, these exercises can usually be completed by many people. The main aim of all the job submission exercises is that people get to know how to specify a job on the Grid, how to help the resource broker find a job match and how to run daily work on the Grid.

In the second afternoon, we first give data management and then information system exercises. Although many people are already familiar with file transfer protocols like FTP, the data management exercises focus on file replication and transfer in a Grid environment. Since many of the replication services follow the web service approach, participants also gain an insight into the field of web services, file cataloguing etc.

The last part of the exercises focus on the information system. Depending on the local installation, people can learn about LDAP technology or relational database technology.

For very advanced students we provide an additional challenge, the so called "Mini project" which requires that students use the available language APIs (mainly Java and C++) to create a small Grid service that optimises data transfer. This exercise is of particular interest since it also introduces the field of network transfer optimisation in a Grid environment.

All exercises and testbed usage instructions are accessible via the tutorial web-site [6]. People can easily do exercises from their office or home, given that they have access to the resources and a valid certificate.

3 Software System

In the following section we provide an overview of the software system that is used for the hands-on exercises. We also give details on what is required to prepare such a tutorial in order to promote the tutorial also for wider use in class room teaching.

In traditional end user tutorials it is often sufficient to install few software packages on each of the students' terminals. However, EDG tutorials require a complex infrastructure (see Figure 2) which includes several servers and it would be too time consuming to install it every time at the tutorial's venue. Thus, it is advisable to have a dedicated infrastructure in place where all the main services are installed and only client side software needs to be added for a specific tutorial session at a remote site.

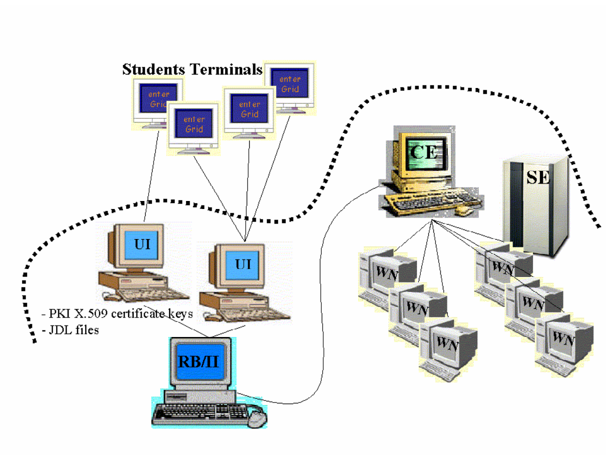

Figure 2. The main Grid components used in a typical EDG tutorial

3.1 A Testbed for Tutorials: GriDis

Current Grid software is typically client-server or peer-to-peer software. The former model is used in EDG and by the underlying Globus Toolkit(tm). Several services interact with each other to allow job execution on a Grid. In EDG we have categorised the required software components and services according to their functionalities and a minimal testbed consists of the following components (compare also with Figure 2):

- **User Interface (UI)**: This refers to all client software components that are actually used by the end users in a tutorial. Typically, the software is installed on the "User Interface machine" which then provides the main access point to the testbed.
- **Resource Broker (RB)**: The Resource Broker is the core of the workload management system and is responsible for deciding where the user job will be executed. All workload management services are supposed to run on a single machine.
- **Computing Element (CE)**: The Computing Element provides the gateway to several **Worker Nodes (WN)** where user jobs are executed.
- **Storage Element (SE)**: The Storage Element provides a Grid storage system for data.
- **Information System (IS)**: The Information System provides basic information about the status of all Grid nodes.

Typically, one machine is required for each of the above components to have a minimal, working Grid infrastructure. This basic structure is not specific to the EDG software system but to all Grid solutions that are based on Globus. Consequently, a minimal set of services and machines needs to be in place in order to provide Grid hands-on exercises. Additional services, like the replica location service, need to be installed but they can be bundled with the existing components described out above. For details of the EDG software system refer to the project web-site [7] and the specific middleware components.

For our tutorials, we set up a basic Dissemination testbed called "GriDis" where all the services as well as the User Interface are provided. The core of the testbed is located in Rome/CNR (UI, RB, CE+WN, 2 SEs, IS and certificate authority) and Geneva/CERN (CE) but it is easily extensible by several machines (additional Computing Elements, Worker Nodes, Storage Elements, etc.) in case the load of the system is increased by a higher number of tutorial participants.

The Information System in GriDis has been configured in a flexible way to make it easy to add or remove sites. All components can easily be reconfigured during a hands-on session to exclude broken or non-responsive services and to allow for redundancy. In addition, user access can be monitored and controlled. This gives us the ability of running more than one event at the time on GriDis. For instance, during a tutorial in Madrid, we opened GriDis for the first time to the public (i.e. not to tutorial participants) during a dissemination event at CERN.

Consequently, GriDis is not only used for EDG Tutorials but also for other dissemination events like demonstrations or the Grid Cafe activity [17], where young students learn about Grids at an early age.

For teaching and demonstration purpose, web portals give easy access to the Grid infrastructure and allow for simple usage of Grid tools via a web browser. On the User Interface machine at GriDis we used the GENIUS portal [15] that provided access to the EDG tools via a web browser. Using GENIUS for teaching purpose has been very successful since the error rate that students have is very much reduced by using predefined tags in the GENIUS system.

3.2 Tutorial Preparation

In order to efficiently prepare for an EDG tutorial, several steps are required that are outlined below. The main emphasis is put on providing a simple solution to host tutorials at remote sites with good network connectivity to the tutorial testbed.

Since one half of the tutorial is class room teaching (see lecture programme in previous section), a conventional lecture theatre with multimedia equipment and Internet connection are required. For each of the lectures, we provide handout material that we give to our participants in the morning of the lecture programme.

Most of the preparation work is dedicated to the hands-on exercises and the preparation of the software system. Experience has shown that the installation of the entire testbed is rather complex and time consuming and thus the basic core services (RB, CE (WN), SE, IS) at GriDis are always in place, Thus, only the User Interface software would need to installed, if necessary, at the location where the tutorial takes place. Like all other EDG components currently, the UI machine should run GNU/Linux RedHat 7.3 and then all EDG specific software components need to be installed and configured. Next, the client tools need to be connected to the GriDis testbed. This requires only a few edits in configuration files.

Once the UI is in place, the basic software infrastructure for the tutorial is ready. Although the UI hosts the basic client software, the UI machine acts as the gateway to which all tutorial participants have to log on from their terminals or laptop computers (see "Students Terminals" in Figure 2). This is to reduce the overhead of installing all client tools on each of the machines.

Services in the EDG software system use the Grid Security Infrastructure that is based on X.509 [19] proxy certificates, i.e. people need to be authorised to use the Grid. One of EDG's main partners provides such certificates that are typically issued a few days before a tutorial can be started and then installed on the user interface machine. In addition, the services need to know about the new users. This is an important point to keep in mind when teaching Grid technologies.

The users are normally registered in a central repository which makes it easy to obtain and include them into the authorisation mechanism of the local site. In addition, GriDis hosts a certificate authority for providing certificates to GriDis users.

4 Experience

Since the first tutorial in summer 2002, we have gained much experience in organising, hosting and running tutorials. Originally, the tutorial programme consisted only of theoretical lectures and was given at the Global Grid Forum in parallel with other Grid tutorials [22]. The initial programme covered four hours but then it was extended to the typical two day events with lectures and practical exercises. It was then further extended and presented at the CERN School of Computing 2002 [2] where for the first time the EDG application testbed was used for the event. Later, we started to run the tutorials on the GriDis testbed.

4.1 Tutorial Events and Users

Up to now (Dec. 2003) the EDG tutorial has been held in the following places (in chronological order):

- Edinburgh (UK) - 2 times (lectures only at GGF)
- Vico Equense (Italy) - 2 times (CERN School of Computing, and GGF Summer School on Grid Computing [21])
- CERN (Geneva, Switzerland) - 4 times
- Torino (Italy)
- Cracow (Poland) - lectures only
- Amsterdam (The Netherlands)
- Lyon (France)
- Tokyo (Japan) - lectures only at GGF
- London (UK)
- Aveiro (Portugal)
- Catania (Italy)
- Seattle (USA) - lectures only at GGF

- Rome (Italy) - 2 times
- Krems (Austria) - CERN School of Computing
- Islamabad (Pakistan)
- Madrid (Spain)
- Athens (Greece)

In total, more than 500 people from academia and industry have been trained. The user feedback we collected from each of the events has contributed to improve both the lecture material as well as the practical exercises. Typically, between 20 and 40 people take part in such an event since this is usually the number of people that can be supervised by two to three tutors in a hands-on session. In special events like the CERN School of Computing and the GGF Grid Summer School between 70 and 80 people participated. Participants usually have different backgrounds in science and engineering and thus it is a challenge to meet all of the user requirements.

Participants usually have different expectations of a tutorial. Some only come to get an overview about existing technologies whereas others want to use or plan to use Grid tools immediately afterwards. Since the certificates issued for the students are usually only valid for the lifetime of a tutorial, people need to request certificates with longer lifetime (usually a year) and then can either join the EDG testbed activities or get a special certificate and account on the dissemination testbed (GriDis). It is important to keep the motivation and thus the momentum of potential users. Therefore, we also give people the possibility to have limited access to a Grid testbed after the tutorial. We consider this as a vital aspect in the field of emerging technologies. We are also aware of the importance of disseminating Grid training at early stages where software tools are not yet as mature as they could be due to the novelty of the technology. Participation to one of the tutorial occasionally resulted in posts in one of the Grid related projects.

4.2 Influence of the Network

One of the main requirements of using Grid technologies and hosting a tutorial is to have good network connectivity to the wide-area network in the order of several Mbps. This is required in order to have a reasonably fast response time between the clients on the user interface and the services on GriDis or the specific testbed used. Usually, wide-area network connections within Europe or the US are fast enough to cope with these requirements.

However, other countries might not have these fast network links and thus it is sometimes difficult to run a conventional tutorial with a user interface at the host site outside Europe and the GriDis testbed in Europe. For our tutorial in Pakistan we installed an entire testbed directly at the hosting institute and thus we were able to run all services locally. This resulted in an additional effort of providing a full installation guide for the entire testbed infrastructure [12] which also allows other people to get experience with installing/configuring and running a small scale testbed. In order to support this, we also provide a DVD with a complete set of all software tools that are required. Note that often the EDG software is installed and configured automatically using LCFG which is usually not available at the tutorial site. Therefore, we needed to provide a full set of manual installation instructions.

4.3 Teaching Grid Computing

Grid computing can be regarded as a special form of distributed computing. It uses several Internet technologies. As for class room teaching, Grid computing is adequate for advanced courses in master's or Ph.D. programmes of Computer Science studies (see Section 5 for further details). Since Grids are particularly aimed at serving computing and data intensive science domains, Grid technologies are also of interest in application science studies.

The lectures of our tutorial programme usually fit into a 10 hour seminar on Grid technologies and this was tried at the University of Perugia in summer 2003 [11]. A Ph.D. seminar at the University of Madrid is in preparation for March 2004 [30]. Furthermore, the tutorial also provides the basic starting point for more focused Grid seminars on specific fields such as workload management, replica management or information systems. Several students projects have already been initiated based on inputs that students gained from attending one of the tutorials.

Since Grid computing is still a rather new research domain, there are several open research fields for potential master or Ph.D. students. Usually, each of the theoretical lectures in our lecture programme covers several aspects where more research can be carried out. Thus, another interesting aspect of our tutorials is to attract new research candidates for the field of Grid computing. As a positive example we stress that our GriDis infrastructure is already used now for practical work of a master student candidate in Rome.

For our tutorial, there are no particular prerequisites for students except the basic knowledge of UNIX and distributed computing. Programming skills and or background in cluster computing are favourable but not required. In general, the course can be applied to graduate students as well as researchers.

4.4 Administration

Running the tutorial programme in a quasi-professional way, requires involvement of several people. In our case, the main tasks consist of keeping the tutorial material up-to-date as well as the maintenance of the GriDis testbed. A dedicated Education&Outreach manager takes care of the overall co-ordination, and a tutorial team with part time participation helps in running the tutorial sessions and updating the material (see Acknowledgement).

As for the testbed, one person is responsible for the main operations of the services but several other people help out in case of emergencies and software upgrades. Usually, not much human intervention is required once the software system is running. For a tutorial session, UNIX user accounts need to be created on the user interface machine as well as all server machines need to be updated with the user certificate information. Although the latter can be automated since there is a general repository with user information, we managed the users manually to quickly intervene when necessary.

4.5 Future Work

Usually, our tutorials run very smoothly once the testbed configuration is done correctly at the hosting site where we run the tutorial. In addition, the lecture programme is well balanced to allow our users to quickly grasp the basics of the software system in order to use it in their daily work. Below are a few points that we are considering in the future in order to improve our tutorials.

- GriDis can be made more specialised and stable with a permanent set of participating sites. We also welcome further sites to join the testbed.

- Although we provide a DVD with all the required software and the installation instructions, the software distribution of the UI can be improved so that it can easily be installed on the fly on several machines and even on portable computers.

- We can also provide a semi-automatic distribution of the software on DVD to install an entire testbed. In this case, a small testbed can easily be installed at each location or campus to provide a small scale internal Grid. Our previous work on automatic software distribution [13] has been very successful and can be applied also to the latest software system provided by EDG.

- A more interactive web-based tutorial could be created whereby users do their exercises through a web portal that gives error messages and instructions on how to improve the input. This can even be done without real access to the Grid but just to preliminary parse possible Grid jobs.

5 Related Work

To the best of our knowledge, our tutorial programme is currently one of the most comprehensive ones in the Grid community since it covers all user aspects including hands-on training. One advantage is also that we often invite developers or people directly involved with the software development process to present their work and provide talks. Here, we give a short overview of related tutorial activities.

NorduGrid: The NorduGrid project provides a tutorial [27] with lectures and practical exercises but is usually not much present at major Grid events and thus does not address a wide user community.

Globus and OGSI: The Globus Alliance provides several user tutorials (mainly lectures) at various Grid events such as GGF meetings or Grid related computer science conferences. A very good tutorial on hands-on training on Globus Toolkit 3 and thus OGSI technology is presented in [29]. The advantage of that tutorial is that it can be done in isolation on a single computer and it does not need a complete testbed to get started. Usually, one needs to install the Globus Toolkit 3 and can then follow the example. In half a day, one can learn to write a simple Grid service. However, lots of background knowledge is required and usually people need to be experienced with web service and Grid technologies to understand the basic concepts.

NPACI: NPACI provides a User Guide (Tutorial) via a web site [28] where users are guided through the account creation process on the NPACI testbed (similar to the EDG User Guide [8]). Furthermore, examples about the usage of Globus, Condor, NWS, Ganglia and APST are given. A separate **Condor tutorial** [4] is directly available on the Condor web site.

Python: Another tutorial offered via a web interface can be found at [20]. It describes how to build a Grid with Python tools.

User tutorials (mainly without hands-on exercises) are often given at major Grid conferences or meetings like GGF, HPDC, or Supercomputing [31].

In addition to software training, several people already provide university courses on Grid technologies [3, 11, 18, 24, 25, 30].

A very good example on how class room teaching is combined with practical exercises on cluster computing can be found in [1]. Due to the complexity of Grid technologies and the distributed nature, many teaching aspects of cluster computing are also valid to Grid computing.

6 Conclusion

Within the EU DataGrid project we have established a comprehensive Grid tutorial aimed at end users that want practical experience with state-of-the-art Grid technology. The tutorial programme has evolved and improved over time based on the feedback of the user community as well as the help of the LHC Grid Computing project (LCG [23]), one of EDG's main partner projects that deploys main parts of the EDG software system.

Our tutorial provides basic training for potential Grid users and thus a fundamental knowledge base on Grid technology. It is not only applicable in class room teaching and in dissemination events and demonstrations but we are also providing fundamental work in promoting Grid technologies in places all over the world (the EDG/LCG tutorial in Pakistan kicked off the very first Grid testbed in central Asia). We are convinced that more training efforts are needed as technology and implementation rapidly change over the next years. Furthermore, Grid technology will need to be included much more into curricula of university programmes in order to provide students the basic requirements for using a challenging and still rather new technology.

Acknowledgements

Special thanks to Mario Reale and Elisabetta Ronchieri for providing the original tutorial hand-out. Thanks to Steve Fisher, Akos Frohner, Bob Jones, Erwin Laure, Charles Loomis, Maite Barroso Lopez, Ricardo Rocha, Massimo Sgaravatto, Antony Wilson, and several others helping the EDG Tutorial team to constantly improve the Grid tutorial material. Thanks to David Cameron for proof-reading this document.

References

[1] Amy Apon, Rajkumar Buyya, Hai Jin, and Jens Mache, Cluster Computing in the Classroom and Integration with Computing Curricula 2001, IEEE Transactions on Education, IEEE Press, USA, 2003. (accepted in 2002 and in print)

[2] CERN School of Computing, http://cern.ch/CSC

[3] Ann Chervenak. CS 599: Introduction to Grid Computing. http://www.isi.edu/ annc/classes/cs599.html

[4] Condor Tutorial: http://www.cs.wisc.edu/condor/tutorials/

[5] EU CrossGrid project: http://www.eu-crossgrid.org

[6] EU DataGrid Tutorial Website: http://cern.ch/edgtutor

[7] EU DataGrid Project: http://www.eu-datagrid.org

[8] EU DataGrid User Guide and Documentation. http://marianne.in2p3.fr/datagrid/documentation/

[9] Karim Djemame. Grid Computing: a New Paradigm for Science and Engineering. http://www.comp.leeds.ac.uk/karim/aiccsa03/

[10] Flavia Donno, Leanne Guy, Mario Reale, Ricardo Rocha, Elisabetta Ronchieri, Massimo Sgaravatto, Heinz Stockinger, Kurt Stockinger, Antony Wilson. EDG Tutorial - Handouts for participants for EDG Release 2.x, Version 3.2, November 2003, https://edms.cern.ch/document/393671/

[11] Flavia Donno. Grid Computing - Dai Clusters a Grid. May 2003, http://pcgrid01.pi.infn.it/perugia/

[12] Flavia Donno, Heinz Stockinger. LCG-1 Tutor Manual Installation Guide, Version 3.0, December 2003. https://edms.cern.ch/document/412774/

[13] Flavia Donno, Andrea Sciaba, Zien Xie. INFNGRID v1.3 Installation Procedure, INFNGRID Note No. 3/2001, March 2001, http://www.pi.infn.it/GRID/dist/GRID_INST_1.3.html

[14] Ian Foster, Carl Kesselman (eds). The Grid: Blueprint for a New Computing Infrastructure. Morgan Kaufman, 1998.

[15] GENUIS Web Portal. https://genius.ct.infn.it/

[16] Grid Dissemination Testbed (GriDis): http://web.datagrid.cnr.it/GriDis/GriDisWP1.html

[17] GridCafe, http://cern.ch/gridcafe/

[18] Geoffrey Fox. Grid Computing Tutorial, http://www.grid2002.org/pgclasssummer03/

[19] Russell Housley, Warwick Ford, Tim Polk, David Solo. Internet X.509 Public Key Infrastructure, RFC 3280, The Internet Society April 2002, http://www.ietf.org/rfc/rfc3280.txt

[20] IBM Developer Work: Build a grid application with Python. http://www-106.ibm.com/developerworks/edu/gr-dw-grpintro-i.html

[21] International Summer School on Grid Computing. Vico Equense, Italy, July 2003, http://www.dma.unina.it/ murli/SummerSchool/

[22] Erwin Laure, Peter Kunszt: EDG Tutorial at Global Grid Forum 5, Edinburgh, July 2002. http://www.gridforum.org/Meetings/ggf5/tutorials.htm

[23] LHC Grid Computing Project (LCG). http://cern.ch/lcg

[24] Peter Lukesch. Lecture on Grid Computing. https://wwwbode.cs.tum.edu/pipermail/gridcomputing/

[25] Gregor von Laszewski. Grid and Ubiquitous Computing at IIT http://www-fp.mcs.anl.gov/ gregor/grid-iit/

[26] Erwin Laure, Heinz Stockinger, Kurt Stockinger. The EU DataGrid Tutorial. CERN School of Computing, Sept. 2003. http://cern.ch/edgtutor/doc/edg-tutorial.pdf

[27] NorduGrid tutorial web site: http://www.nordugrid.org/tutorial/

[28] NPACI User Guide - Tutorial. http://npacigrid.npaci.edu/tutorial.html

[29] Borja Sotomayor. The Globus Toolkit 3 Programmer's Tutorial. http://www.casa-sotomayor.net/gt3-tutorial/

[30] Heinz Stockinger. Grid Computing for Data Intensive Applications. http://cern.ch/hst/programme-grid-seminar.html

[31] Tutorials at Supercomputing 2003: http://www.sc-conference.org/sc2003/tech_tutorials.php