Nordic Testbed for Wide Area Computing and Data Handling

2001-12-20

NORDIC TESTBED FOR WIDE AREA COMPUTING AND DATA HANDLING (NORDUGRID)

Mid-term Report

Introduction

The purpose of the project is to create the grid computing infrastructure in Nordic countries. Project participants include universities and research centers in Denmark, Sweden, Finland and Norway. The active phase of the project started in May 2001, and involves the Niels Bohr Institute (Denmark), Lund and Uppsala Universities (Sweden), University of Oslo (Norway) and Helsinki Institute of Physics (Finland). From the very beginning, the NorduGrid testbed became an integral part of the EU DataGrid project, initiated at CERN and aimed at creation of the computing infrastructure for the future high-energy physics experiments. This report overviews the status of the NorduGrid project as of December 2001.

1. Organization

Due to the considerable geographical spread of the project member sites, the organization and management heavily relies on the Internet and teleconferences. The project Web site *http://www.nordugrid.org* opened in December 2000, and is regularly updated with the most detailed information.

1.1. Responsibilities

1.1.1. Steering Panel

Prior to the project's start, the steering panel was appointed to provide adequate management. It consists of the coordinator, John Renner Hansen, representing Niels Bohr Institute, Denmark; Tord Ekelöf, Paula Eerola and Sten Hellman, representing Uppsala, Lund and Stockholm Universities of Sweden; Farid Ould-Saada from the University of Oslo, Norway, and Matti Heikkurinen, representing Helsinki Institute of Physics, Finland. The group holds regular phone meetings.

1.1.2. New positions

Three new postdoctoral positions, funded by the project, were created. Each employed researcher carries responsibility for one of the NorduGrid nodes: in Lund, Uppsala and Oslo, and provides support and assistance to other sites in the area. Positions were announced in December 2000, and chosen by the steering panel were:

Name	Location	Start
Balázs Kónya	Lund	May 1st, 2001
Mattias Ellert	Uppsala	May 15, 2001
Aleksandr Konstantinov	Oslo	June 20, 2001

Table 1. New positions

These researchers carry out most of the project's activities, including set up and maintenance of the testbed sites, evaluation, necessary development, and user support.

1.1.3. Technical Working Group

For the fast evaluation of existing demands and problems, a technical working group, consisting of the three abovementioned postdocs and people, responsible for the cooperation with EU DataGrid (see Section 5): Anders Wäänänen (Testbed and Demonstrators Work Package), and Oxana Smirnova (HEP Applications Work Package). The group meet bi-monthly and discuss status of the project, relations with EU DataGrid, various problems, and provides an immediate short-term planning.

The first (introductory) meeting took place at NBI, on June 20-22, 2001. Present were: B.K., M.E., A.K., A.W., J.R.H., F.O.-S.. Several important decisions were made:

- Web-site structure
- Creation of the NorduGrid mailing lists: one for general, and one for the technical discussions
- Details on the Certificate Authority setup and related procedures (see Section 3)

- Installation of three dedicated clusters (see Section 2)
- Nearest tasks (ca. half a year)

The second meeting was held on August 21, 2001, in Lund. The meeting was the follow-up, and the discussed issues covered acquired experience, the resource specification for the DataGrid Testbed1 (see Section 5), user policy, Globus toolkit upgrade and future steps.

The third meeting was held in Oslo, in the framework of the 2nd NorduGrid Workshop, on October 31 and November 1, 2001. The main issues discussed were: the implementation of the EU DataGrid middleware, and the preparation of the Globus installation hands-on tutorial.

1.2. General Meetings

Apart of the steering panel and the technical working group meetings, a few general workshops are foreseen.

The inaugural 1st NorduGrid Workshop took place in Lund, February 5-6, 2001. Present were coordinators of Grid-related activities in participating institutions, and researchers, otherwise involved in the project. The purpose of the workshop was the general presentation of the project and its planned activities, review of the situation in the Nordic countries, and discussion of organizational issues. During the Workshop, two parallel sections were working: one covering the testbed installation issues (chaired by A.W.), and another - discussing applications (chair: O.S.). A presentation of the business aspects and industrial applications for the grid was made by M.H..

The 2nd Workshop took place in Oslo, November 1 to 2, 2001. Major issues were:

- First hands-on experience with installation and usage of the hardware and middleware
- Demonstrations of Globus™ installation and the Grid portal
- Partnership with software companies and computing centers
- Grid applications, ATLAS Data Challenges

2. Resources

The whole project infrastructure relies on the existing networks in the Nordic countries. The connectivity chart for the participating sites is shown in Figure 1

From the project start, part of existing computing resources was assigned for the test installations. The LSCF cluster at NBI (ca. 30 heterogeneous processors) was used for initial middleware and resource management tests. In Lund, a mini-farm (2 PC's) was set up at the Particle Physics department, and a stand-alone machine was configured for Grid tests at the Cosmic and Subatomic Physics department. In Uppsala University and Helsinki Institute of Physics, a stand-alone PCs were originally used for initial tests.

In addition to existing resources, five new dedicated clusters were set up: in Lund, Uppsala, Oslo, Bergen and Copenhagen. Lund and Uppsala clusters upgraded and replaced previously used local test facilities.

Figure 1. Connectivity map of the NorduGrid participating sites

Computing resources are fairly heterogeneous, hardware- and software-wise. Most of the processors are various types of Intel Pentium. Operating systems are various flavors of Linux: different distributions of Red Hat, Mandrake, Slackware and Debian. On test clusters, resource management is currently performed via OpenPBS.

Grid middleware was chosen to be uniform, and test sites are set up with Globus toolkit version 2b.

Detailed specifications for each cluster are shown below.

2.1. Uppsala cluster

Name	grid.tsl.uu.se	**Contact person**	Mattias Ellert	grid.support@tsl.uu.se	
Address	Uppsala University, Department of Radiation Sciences, Uppsala, Sweden				
HARDWARE					
Nodes	*Quantity*	*CPU*	*RAM*	*Disks*	*Other notes*
	1	P-III 866 Mhz	512 MB	40 GB + 60 GB	Gatekeeper
	2	2 × P-III 866 Mhz	512 MB	40 GB	Computing nodes
Network	100 Mbit/s				
Mass storage	None				
SOFTWARE					
OS	RedHat Linux 7.1				
	Kernel 2.4.9-13 on the gatekeeper, kernel 2.4.9-13smp on the computing nodes glibc 2.2.4-19				
Resource manager	OpenPBS 2.3.2				
File system	/home and /usr/local shared through NFS				
Databases	None				
MIDDLEWARE					
Gatekeeper	/O=Grid/O=NorduGrid/CN=grid.tsl.uu.se/jobmanager, /O=Grid/O=NorduGrid/CN=grid.tsl.uu.se/jobmanager-pbs				
Globus	Globus 2b21 (EDG)				
NETWORK					
Configuration	Private network				
COMMENTS					
The computers are on a private network. Only the gatekeeper is directly accessible from the outside. The outside is however directly accessible also from the computing nodes. Following cells are mounted using OpenAFS 1.2.2: /afs/cern.ch, /afs/desy.de					

Table 2. Uppsala cluster specifications

2.2. Lund cluster

Name	Lund Grid Cluster	**Contact person**	Balázs Kónya	grid.support@grid.quark.lu.se	
Address	Elementary Particle Physics Department of Lund University				
HARDWARE					
Nodes	*Quantity*	*CPU*	*RAM*	*Disks*	*Other notes*
	1	PIII (Coppermine) 1GHz 256KB cache	512 MB	40 & 60 GB	Front-end machine, Powerware 5115 UPS
	2	Dual PIII (Coppermine) 1GHz 256KB cache	512 MB	30 GB	Dual processor computing nodes
	1	PIII (Coppermine) 1GHz 256KB cache	256 MB	16 GB	Single processor computing node
Network	100Mb/s private network with a 3Com OfficeConnect 16 100MB Switch				
Mass storage					
SOFTWARE					
OS	Linux Mandrake 8.0 distribution, kernel-2.4.3				
Resource manager	OpenPBS 2.3.12				
File system	All the nodes have local disks and share the NFS mounted /scratch and /home area of the front-end machine				
Databases					
MIDDLEWARE					
Gatekeeper	grid.quark.lu.se 2119 port				
Globus	Globus version 2beta (Globus)				
NETWORK					
Configuration	All the nodes are on a closed private network behind the front-end machine.				
COMMENTS					
The grid cluster of the Elementary Particle Physics Department of Lund University, dedicated to the NorduGrid project, contains six Intel Pentium III 1GHz processors with 256 MB RAM per processors. The cluster consists of four Linux based Pcs, two of them are dual-processor machines. The cluster is made up of a front-end machine (single-processor) and three computing nodes with 5 available Pentium III 1GHz processors (1 single and 2 dual nodes). The computing nodes are connected to the front-end machine using a private network, which means that the nodes can only be accessed through the front-end computer. The front-end machine (grid.quark.lu.se) runs PBS as the local resource management system. The front-end node of the cluster is dedicated for code development (code editing, compilation, etc.), while the back-end nodes (node1, node2, node3) are used only for code executions					

Table 3. Lund cluster specifications

2.3. Oslo cluster

Name	Oslo University NorduGrid gatekeeper	**Contact person**	Aleksandr Konstantinov	grid.sysadmin@fys.uio.no grid.support@fys.uio.no	
Address	University of Oslo, Department of Physics				
HARDWARE					
Nodes	*Quantity*	*CPU*	*RAM*	*Disks*	*Other notes*
	1	2 x Intel PIII 1 GHz	256MB	39266MB	computing node
	1	Intel PIII 1 GHz	256MB	39266MB	computing node
	1	Intel PIII 870 MHz	128MB	2 x 41174MB	gatekeeper
Network	100Mbps ethernet cards EtherExpress Pro100				
Mass storage	NA				
SOFTWARE					
OS	Gatekeeper - distribution: RedHat 7.1 , kernel: 2.4.2-2 (i686) , libc: 2.2.2				
	Nodes - distribution: Slackware 7.1 , kernel: 2.2.19 (i686) (SMP & UP) , libc: 2.2.3				
Resource manager	OpenPBS 2.3.12				
File system	ext2fs				
Databases	NA				
MIDDLEWARE					
Gatekeeper	grid.uio.no 2119				
Globus	Globus 2b21 (EDG)				
NETWORK					
Configuration	Nodes are situated on private physically isolated network with 100Mbps connection through Allied Telesyn FS708 network switch.				

Table 4. Oslo cluster specifications

2.4. Bergen cluster

Name	Bergen University NorduGrid gatekeeper	**Contact person**	Aleksandr Konstantinov	grid.sysadmin@fys.uio.no grid.support@fys.uio.no	
Address	University of Bergen, Department of Physics				
HARDWARE					
Nodes	*Quantity*	*CPU*	*RAM*	*Disks*	*Other notes*
	2	2 x Intel PIII 1 GHz	1024MB	41174MB	computing node
	1	Intel Celeron 800 MHz	512MB	2 x 41174MB	gatekeeper
Network	100Mbps ethernet cards EtherExpress Pro100				
Mass storage	NA				
SOFTWARE					
OS	Gatekeeper - distribution: RedHat 7.1 , kernel: 2.4.7-10 (i686) , libc: 2.2.4				
	Nodes - distribution: Slackware 7.1 , kernel: 2.4.14 (i686) (SMP) , libc: 2.2.3				
Resource manager	OpenPBS 2.3.12				
File system	Nodes: ext2fs, gatekeeper: ext3fs				
Databases	NA				
MIDDLEWARE					
Gatekeeper	grid.uio.no 2119				
Globus	Globus 2b21 (EDG)				
NETWORK					
Configuration	Nodes are situated on private physically isolated network with 100Mbps connection through network switch.				

Table 5. Bergen cluster specifications

2.5. Copenhagen cluster

Name	Niels Bohr Institute NorduGrid gatekeeper	**Contact person**	Anders Wäänänen	grid.sysadmin@nbi.dk grid.support@nbi.dk	
Address	Niels Bohr Institute, Copenhagen				
HARDWARE					
Nodes	*Quantity*	*CPU*	*RAM*	*Disks*	*Other notes*
	3	2 x AMD 1.2 GHz	512MB	41174MB	Computing node
	1	Intel 1GHz	512MB	2 x 41174MB	gatekeeper
Network	100Mbps ethernet cards EtherExpress Pro100				
Mass storage	NA				
SOFTWARE					
OS	Gatekeeper - distribution: RedHat 7.2, kernel: 2.4.9-13 (i686), libc: 2.2.4				
	Nodes - distribution: RedHat 7.2, kernel: 2.4.9 -13 (i686) (SMP), libc: 2.2.4				
Resource manager	OpenPBS 2.3.12				
File system	Ext3, nfs				
Databases	NA				
MIDDLEWARE					
Gatekeeper	grid.nbi.dk 2119				
Globus	Globus 2b21 (EDG)				
NETWORK					
Configuration	Nodes are situated on private physically isolated network with 100Mbps connection through network switch.				

Table 6. NBI cluster specifications

3. Grid Services

To provide proper functionality of a computational grid, several services should be enabled. Installation and maintenance of such services is the essential part of a testbed set-up. The present status of NorduGrid services is described below.

3.1. Certification Authority

User authentication is one of the key issues in a Grid environment. Globus toolkit uses personal certificates, issued by a recognized certification authority, to identify each user. The NorduGrid Certification Authority (CA) is set up at NBI. It provides X.509 certificates for identification and authentication purposes. The scope is limited to people from the Nordic countries involved in Grid-related projects: primarily the NorduGrid and EU DataGrid, as well as DKGRID (Denmark). Contrary to most Certificate Authorities worldwide, the NorduGrid one is not a national, but a transnational virtual organization.

The certificates are meant to be used with the Globus toolkit, to provide user authentication. They are recognized not only by the sites, participating in the NorduGrid project, but also by the EU DataGrid, in the framework of the DataGrid Testbed Work package.

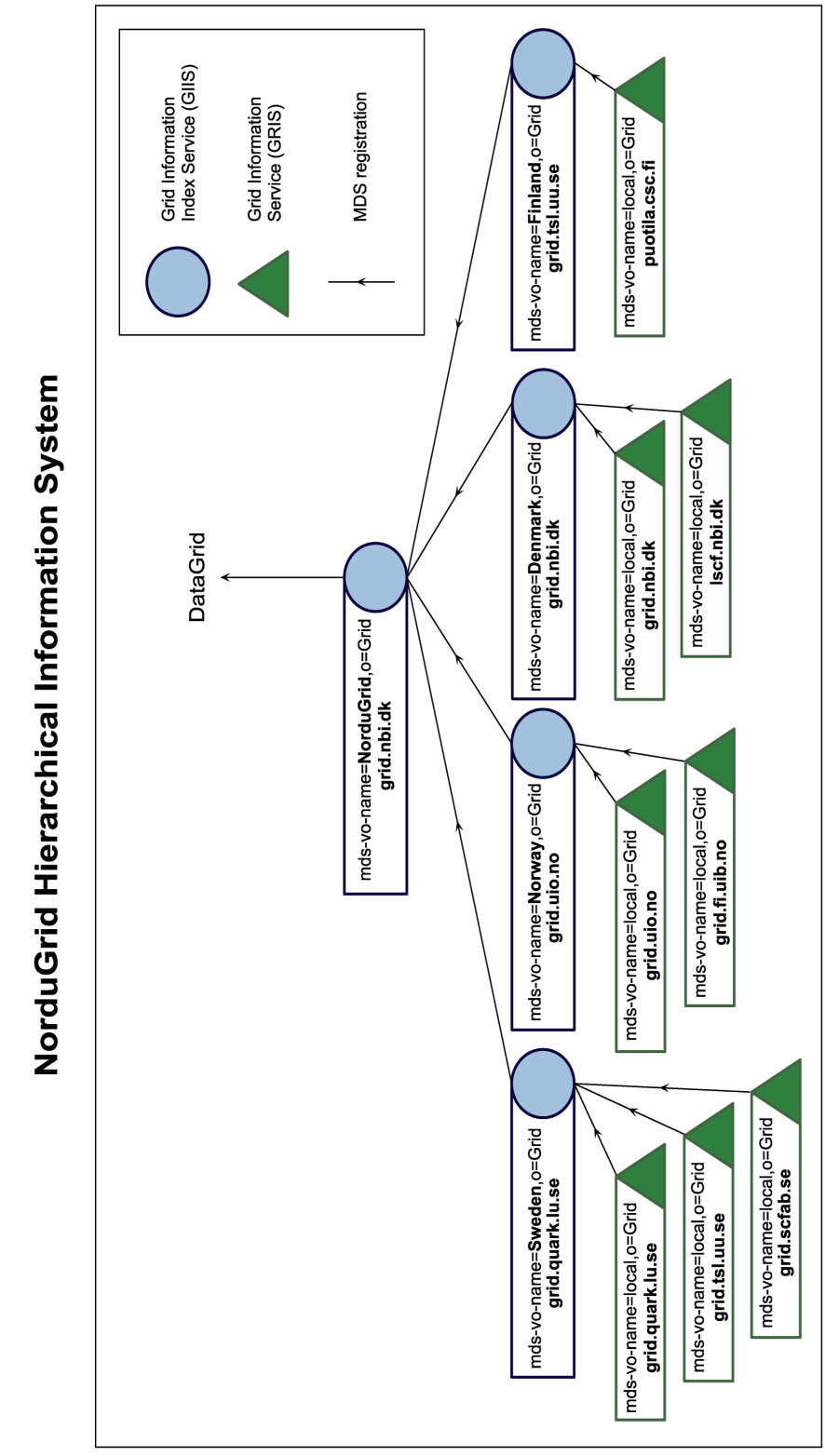

Figure 2. NorduGrid information system

3.2. Information Service

The Globus toolkit provides means of querying resources on a computational grid for their current configuration, capabilities, and status. Such an information is essential for proper distribution of workload. The corresponding database server is running at NBI (ldap://giis.nordugrid.org:2135), providing information on known resources. The browsable index is accessible via WWW, and the schematic layout of the information tree is shown in Figure 2.

The NorduGrid project has adopted a common naming convention for identifying its resources and users. The agreed namespace represents the resources of the project as part of a virtual organization (O=NorduGrid). The distinguished name of a resource has the form of "/O=Grid/O=NorduGrid/CN=grid.quark.lu.se", where the CN field is the name of the computing resource. A NorduGrid user is identified with the "O=Grid/O=NorduGrid/OU=domain.name/CN=User Name" string, here the OU field is the domain name used by his home institute and the CN field contains his real name.

3.3. Virtual Organization

The NorduGrid Virtual Organization (VO) is the user and service management system, which consists of the user and service database and a set of tools for database and grid-mapfile management. In particular, the NorduGrid users gain access to the NorduGrid resources by being added to the NorduGrid VO.

In a Grid environment, users usually don't have login accounts with passwords for the computing resources they want to use, rather they hold a certificate issued by a Certificate Authority. This certificate authenticates them to the required resources. Authentication, however, does not mean automatic access to the resource. Access control for the computing resources (authorization) is an issue of a local policy, in the Grid environment it is done by mapping the accepted set of user certificates to local user accounts.

The NorduGrid VO maintains a list of accepted users who are authorized to use the NorduGrid resources. The VO tools provide an automatic method for the NorduGrid sites to easily maintain the NorduGrid VO user to local Unix account mappings.

The VO is intended to maintain not only the user data, but all the services running on the NorduGrid, which require secure authentication and authorization, that is, they possess certificates. The database will be used to generate grid-mapfiles for these grid services, too.

The VO database server runs from Lund (ldap://grid-vo.nordugrid.org:389), and can be accessed via the project's Web site.

3.3.1. NorduGrid User Policy

Access to the NorduGrid resources can be granted to the following users:

1. Members of the NorduGrid Virtual Organization, i.e., people
    - affiliated with one of the NorduGrid partners
    - agreed to the Acceptable Use Policy document (available on the Web site)
2. Members of other Virtual Organizations, accepted by the NorduGrid Steering Committee

Decision of acceptance is taken on case-by-case basis. Applications should be submitted to the NorduGrid contact persons.

3.4. GDMP and Replica Catalog

Globus Replica Catalog consists of a database server, storing information about data files situated on different storage elements (SE). This information is managed using Globus API and tools.

GDMP (Grid Data Mirroring Package) is the tool for replication of data files over network developed in collaboration between the EU DataGrid project and the Particle Physics Data Grid Project (PPDG). The purpose of GDMP is to replicate data between different storage elements (SE). GDMP uses some additional attributes, not available in Replica Catalog.

To provide Globus-compatible secure authentication to the Replica server, modification patches were made available at the NorduGrid sites. The NorduGrid Replica Catalog (server in Oslo, ldap://grid.uio.no:389) can be browsed via the project's Web site.

3.5. GridFTP

The GridFTP protocol was introduced by the Globus project to provide a secure, reliable, efficient and high performance data transfer mechanism over high-bandwidth wide area networks, which is a key component of any kind of Grid infrastructure. It contains extensions to the standard highly popular FTP protocol, in order to meet the requirements of high performance wide area data movement. The supported features are:

Grid Security Infrastructure

partial file transfer

third-party (from server to server) transfer

automatic negotiation of TCP buffer sizes

parallel (multi-threaded) data transfer.

The GridFTP servers are implemented at all the NorduGrid sites. A study done by the NorduGrid confirmed the superiority of this protocol over other available tools, particularly when used over high-capacity networks, like the one provided by the NorduNet.

4. Applications

The present status of the Globus toolkit allows for simple tests only, checking the connectivity and basic functionality. Following the Globus installation at all the sites, the most trivial tests were made, and inter-connectivity was successfully checked between NBI, Lund, Uppsala, Oslo, Bergen, Stockholm and Helsinki.

For further tests, more advanced applications, relying on realistic physics analysis cases, are being prepared. Starting year 2002, it is foreseen to use NorduGrid resources in the framework of the ATLAS Data Challenges, which constitute extensive tests of the ATLAS Collaboration software and its readiness for the data taking.

Below is the list of tasks, preliminary tested for a local batch submission, and for a rudimentary remote submission via Globus.

4.1. Boson source elongation study

Task description	Study of the pion source elongation in Z decays. Uses Jetset/Pythia to generate e+e- hadronic events.
People	Raluca Muresan (NBI), Oxana Smirnova (Lund)
Executable	File size: 1.8 MB, occupied memory: 16 MB
Input	ASCII input cards (40 B)
Output	Binary HBOOK file (0.5 to 1 MB)
Specific requirements	CERNLIB and Jetset libraries needed for compilation

Table 7. Application: boson source elongation study

4.2. Fragmentation functions study

Task description	Study of hadronization corrections to the helicity components of the fragmentation function in hadronic decays of Z boson. Uses Jetset/Pythia to generate e+e- hadronic events.
People	Oxana Smirnova, Christina Zacharatou Jarlskog (Lund)
Executable	File size: 1.2 MB, occupied memory: 2 MB
Input	ASCII input cards (40 B)
Output	Binary HBOOK file (12 KB)
Specific requirements	CERNLIB and Jetset libraries needed for compilation

Table 8. Application: fragmentation functions study

4.3. Kaon production in the string model

Task description	Study of identical kaons production in the Lund string model. Uses Jetset/Pythia to generate e+e- hadronic events.
People	Oxana Smirnova (Lund)
Executable	File size: 0.8 MB, occupied memory: 2 MB
Input	-
Output	Binary HBOOK file (360 KB)
Specific requirements	CERNLIB and Jetset libraries needed for compilation

Table 9. Application: kaon production in the string model

4.4. Event generation and simulation for ATLAS

Task description	Monte-Carlo generation (PYTHIA) + ATLFAST (for ATLAS project)
People	Børge Kile Gjelsten (Oslo)
Executable	File size: 16 MB, occupied memory: 6 MB
Input	Text file
Output	Text and HBOOK files
Specific requirements	CERNLIB and PDFLIB libraries needed for compilation Statically linked binary can be produced

Table 10. Application: event generation and simulation for ATLAS

5. Participation in DataGrid

The NorduGrid project participates in the EU DataGrid activities along two directions:

- being part of the DataGrid Testbed (coordinator: A.Wäänänen)
- development of applications (coordinator: O.Smirnova)

All the NorduGrid sites successfully took part in the Testbed0 (June 2001), main goal of which was Globus toolkit and basic services installation. The NorduGrid is an integral part of the Testbed, and the NorduGrid CA is one of 11 officially recognized by DataGrid.

The next phase, Testbed1, started in September 2001, and involves installation of the first release of the DataGrid middleware, together with validation and execution of test use-cases. NorduGrid participates this phase both with manpower (A.Wäänänen – testbed integration, and O.Smirnova – validation), and with hardware resources, installing and testing necessary software with purpose to provide computing resources to the EU DataGrid Testbed1.

Participation in the Applications Work Package of DataGrid proceeds via the ATLAS experiment, which distributes the so-called ATLAS Toolkit, containing three physics use-cases. Installation of the toolkit was successfully done at the Uppsala and Oslo sites. In future, it is foreseen to set up full-scale runtime environments for experiments like ATLAS and ALICE, at some of the NorduGrid sites.

As a part of the DataGrid, the NorduGrid sites are meant to be used not only for physics applications, but also for other tasks, like, e.g., biology. In the framework of Testbed0, there were contacts established with the biologists in Lyon, France, testing job submission and file transfer between Lyon and NorduGrid sites.

To enable close cooperation with the DataGrid, representatives of the NorduGrid regularly attend corresponding meetings and Workshops. Presentations of the NorduGrid activities were made at the First DataGrid workshop in Amsterdam (March 2001), and the Second DataGrid workshop in Oxford (July 2001).

6. Summary and Outlook

During the covered period, much progress have been made, and all the planned milestones met successfully. The basic grid infrastructure and services are set up and being constantly upgraded to meet the needs of the testbed. The NorduGrid is the unique geographically distributed international testbed, providing maximum possible functionality with available toolkits and middleware. An invaluable experience is being acquired and documented on the project's Web site.

To achieve a full-scale functionality, the NorduGrid project relies on the development by the EU DataGrid. With the first release of the DataGrid tools in December 2001, a significant step towards implementation of a fully functional distributed computing network.

For the further development of the NorduGrid testbed, the following major steps are foreseen:

- Installation and validation of the DataGrid tools (including Globus 2 release)
- Participation with the resources in the EU DataGrid Testbed1
- Inclusion of more test clusters, both from existing resources and newly acquired ones
- Adaptation of more use-cases to be used as test applications, physics and other sciences alike; participation in the ATLAS Data Challenges is of particular value
- Education of future users: hands-on tutorial planned beginning 2002

Further information can be found on: *http://www.nordugrid.org/papers.html*