

NorduGrid

**From World Wide Web to World Wide Grid
- creating a Nordic Testbed for
Wide Area Computing and Data Handling**



www.nordugrid.org





The aim of this brochure is to describe the NorduGrid research project and its achievements. The NorduGrid research project was funded by NorduNet2 during the years 2001-2002.

The project coordinator is Prof. John Renner Hansen, NBI. The NorduGrid Steering Group consists of Prof. P. Eerola (Lund), Prof. T. Ekelöf (Uppsala), Prof. J. Renner Hansen (NBI), Docent S. Hellman (Stockholm), Assoc. Prof. F. Ould-Saada (Oslo), Programme Director J. Tuominiemi (HIP).

The NorduGrid Technical Working Group consists of Dr. M. Ellert (Uppsala), Dr. A. Konstantinov (Oslo), Dr. B. Kónya (Lund), Dr. O. Smirnova (Lund), Dr. A. Wäänänen (NBI).

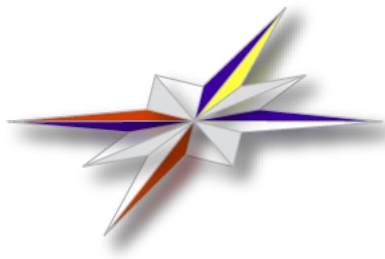
For more information, contact us at contact@nordugrid.org or visit our homepage at <http://www.nordugrid.org>

Copenhagen, October 2002.



www.nordugrid.org





NorduGrid objectives and participants

The Nordic Testbed for Wide Area Computing And Data Handling project (Nordugrid) is a part of the Nordunet2 programme, aimed at developing networked applications with extensive usage of modern utilities and tools in the Nordic countries.

Objective

The aim is to establish an inter-Nordic test bed facility for implementation of wide area computing and data handling. The facility will provide the infrastructure for interdisciplinary feasibility studies of the Grid-like computer structures. The project shall collect and document experience, as an input to the decision process on the future computer infrastructure strategy for sciences with distributed PByte storage requirements and processing power in the order of multi-teraflop.

What is the Grid?

The Grid is a technology to share and access seamlessly computing resources that are not subject to a centralized control. The Grid will be the future infrastructure of computing and data management. The computing resources are connected together through a layer of software called the middleware, which uses standard, open, general purpose protocols and interfaces. This middleware forms the glue binding the resources into a virtual system.

In the same way as the World Wide Web gives us access to information, the World Wide Grid will give us access to computing capacity and data storage in the future.

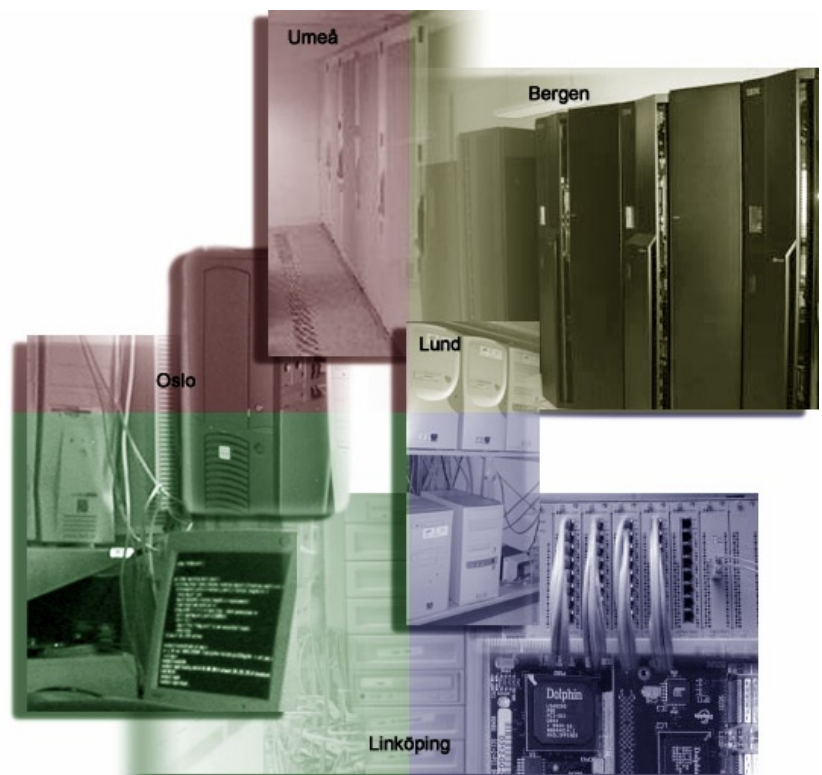
Participants

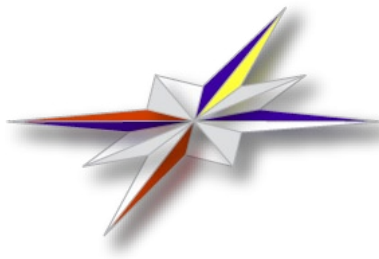
DK Research Center COM
DK DIKU
DK Niels Bohr Institute

SE Lund University
SE Uppsala University
SE Stockholm University
SE Royal Institute of Technology

NO Oslo University
NO Bergen University

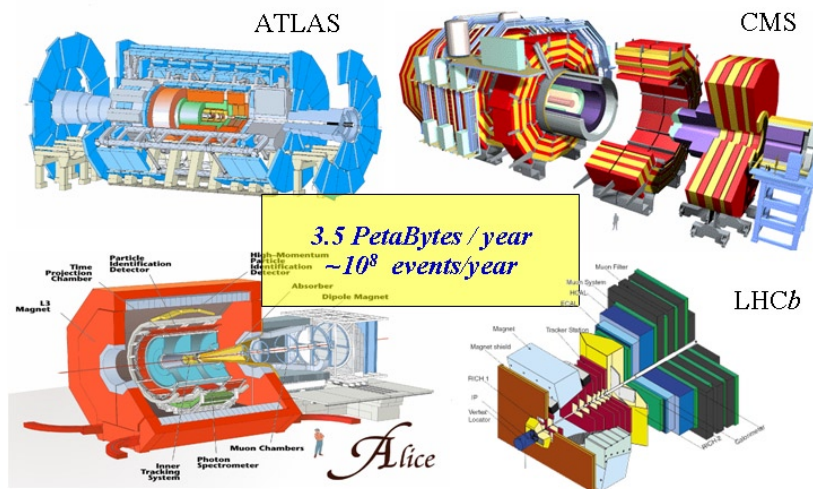
FI Helsinki Institute of Physics





Why do we need Grid in the Nordic countries?

The Grid will play a decisive rôle in the future development of computer and data management in society. The implied future access to vast data bases for scientific, technological, geological, meteorological, economic, demographic, literary, artistic etc. data and the use of enormous computer capacity to analyse and synthesize these data will radically enhance the possibilities for scientific and technological research, industrial and commercial management, cultural and political activities and so on. It is therefore of prime interest for the Nordic countries to participate actively at the international frontier of this technical development. Below is given, as an example of this development, the particular application of the Grid to High Energy Physics research that is presently used to drive the development of NorduGrid.



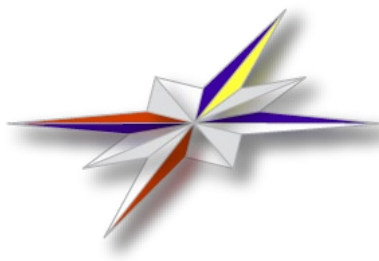
Ever increasing amounts of data: an example

The European High Energy Physics community is in the process of construction and deployment of the Large Hadron Collider (LHC) - the world's biggest accelerator, being built at the European Particle Physics Laboratory (CERN) in Geneva.

Challenges to be faced by physicists are unprecedented. Four experiments will be constructed at the LHC to observe events produced in proton-proton and heavy ion collisions. Data collected by these experiments will allow for exploration of new frontiers of the fundamental laws of nature, such as the Higgs mechanism with possible discovery of the Higgs boson, CP-violation in B-meson decays, supersymmetry, extra dimensions, mini black-holes, and so on.

One of the greatest challenges of the LHC project will be the acquisition and analysis of the data. When, after a few years of operation, the accelerator runs at its design luminosity, each detector will observe bunch collisions at a rate of 40 million per second. A set of filter algorithms, implemented in hardware and on state-of-the-art programmable processors, aims to reduce the event rate to less than 1000 events per second for final storage and analysis. The equivalent data volume is between 100 MByte/sec and 1 GByte/sec. Each experiment is expected to collect 1 PByte of raw data per year. The two LHC general purpose experiments, ATLAS and CMS, have each more than 150 participating institutes distributed all over the world. 2000 physicists per experiment contribute to the development of hardware and software and they expect to have almost instantaneous access to the data and to a set of up-to-date analysis tools.





Nordugrid achievements

Project description

Nordugrid is a Grid research project which:

- develops openly available middleware (Nordugrid Toolkit),
- operates a production quality Grid Testbed,
- pursues basic research on the Grid computing and surveys current Grid technologies,
- exposes the infrastructure to end-users in different scientific communities.

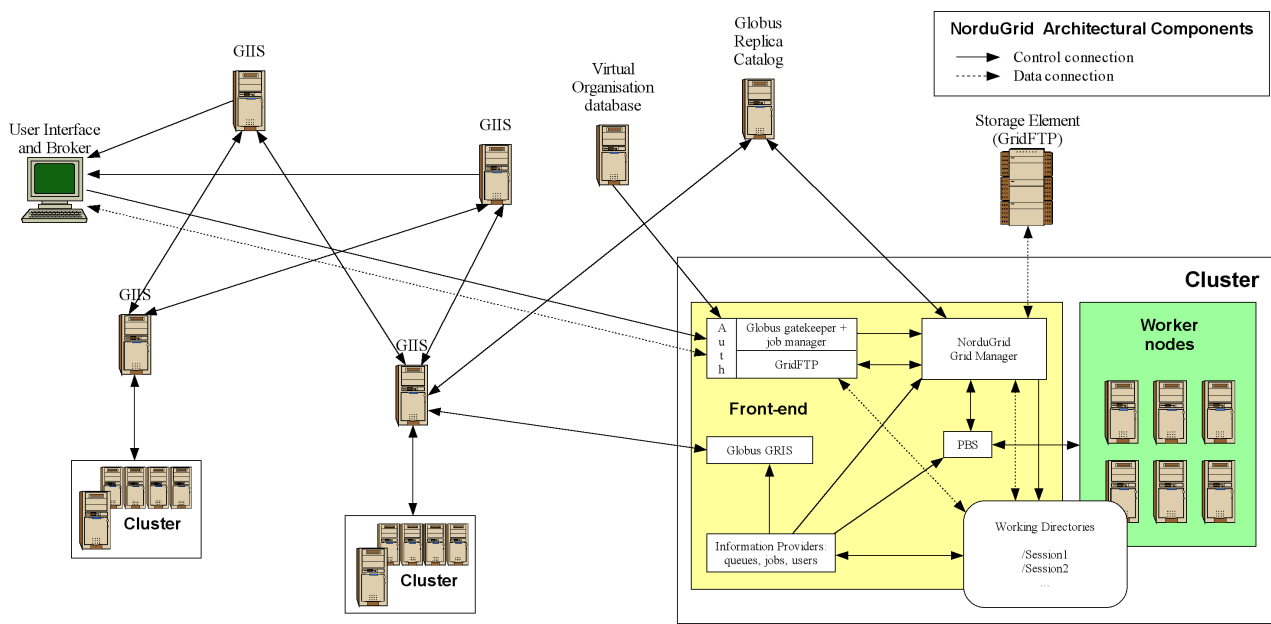
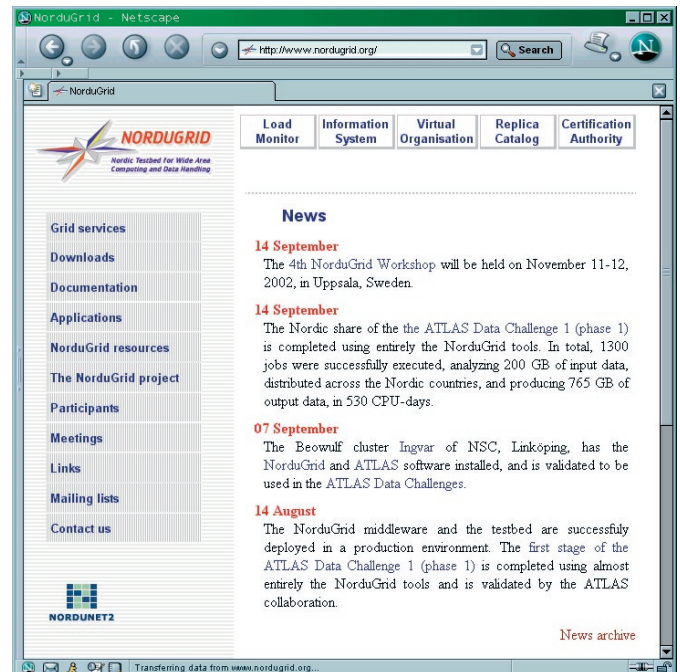
The Nordugrid Testbed consists of 10 sites in the Nordic countries, with about 200 CPUs and 4 Tb of disk storage capacity in total.

The architecture

The Nordugrid architecture consists of the following elements:

- Information System, including Nordugrid Information Model
- Grid Manager
- User Interface and Broker
- Storage Element
- Replica Catalog

The Nordugrid architecture is described in more detail in the documents available from the project's homepage at <http://www.nordugrid.org/documents>

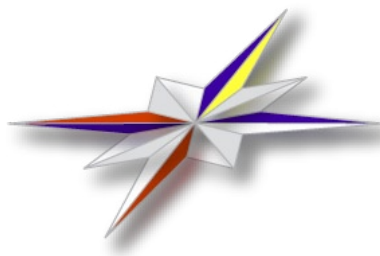


October 2002



www.nordugrid.org





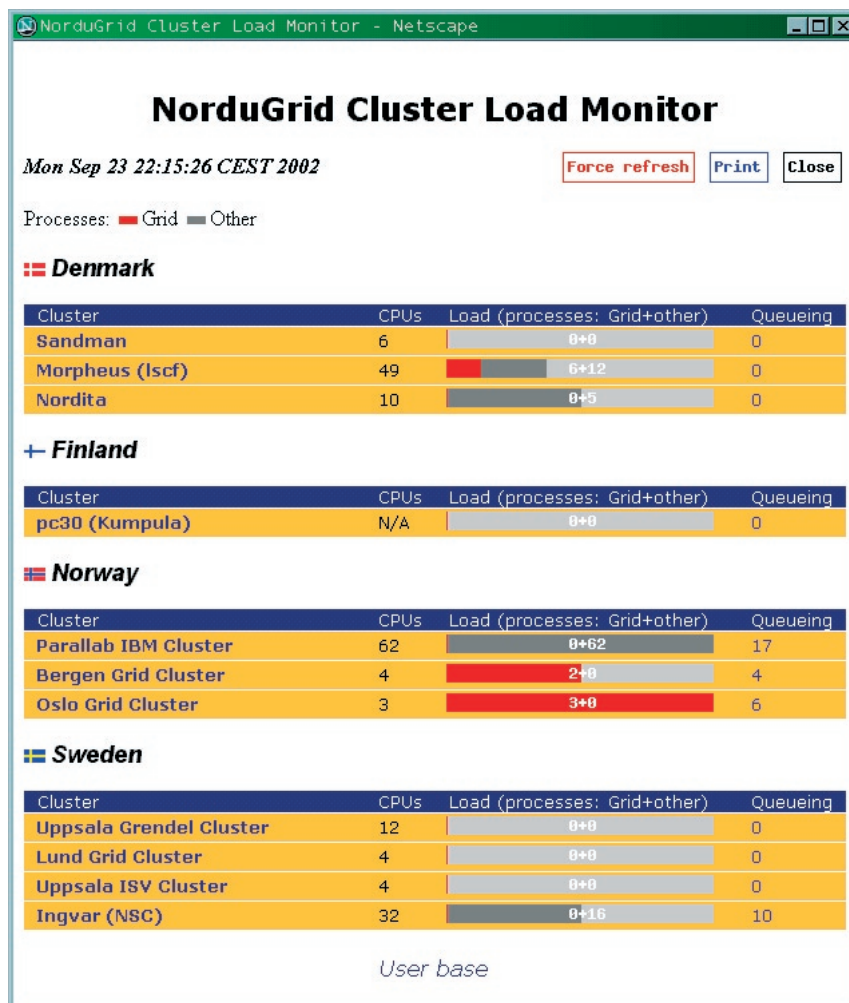
Nordugrid Toolkit

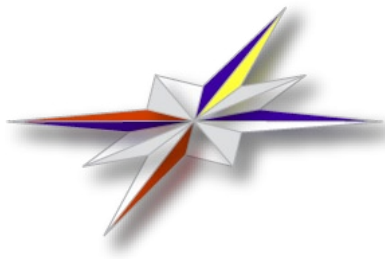
The Nordugrid toolkit is a set of Globus-based Grid middleware, developed to meet the basic requirements of distributed computing and data handling. The toolkit supports a broad variety of Linux platforms (RedHat, Mandrake, Slackware, Debian) and it is freely available from the Nordugrid site at <http://www.nordugrid.org>.

Its most important components are:

- the MDS-based Information System, providing reliable dynamic information about the status of the services, clusters, jobs and users,
- the Grid Manager, residing at a gatekeeper and performing job pre- and post-processing, as well as job control and submission to the local resources,
- the User Interface, responsible for the user's interaction with the Grid. The User Interface contains a built-in resource broker,
- extended resource specification language (XRSL) for formulating job requests,
- Load Monitor web interface to the Nordugrid Testbed.

The software package is easily installable on top of an existing Globus installation and comes with detailed documentation.





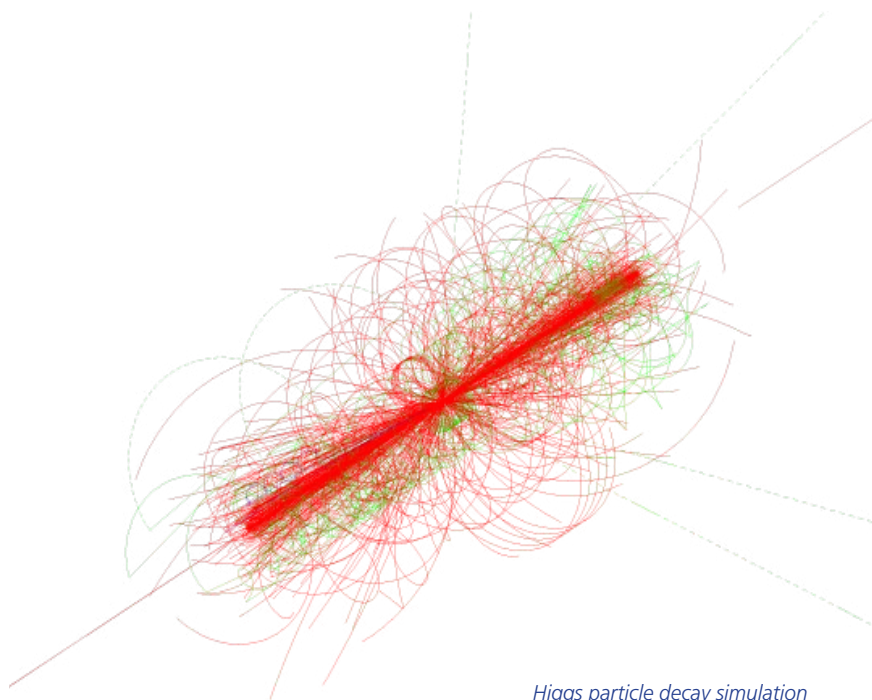
Nordugrid in production

During the summer of 2002, the Nordugrid testbed and middleware were successfully deployed in a production environment. Nordic physicists participating in the ATLAS experiment at the LHC used the Nordugrid toolkit to simulate a large amount of data.

The ATLAS Data Challenges

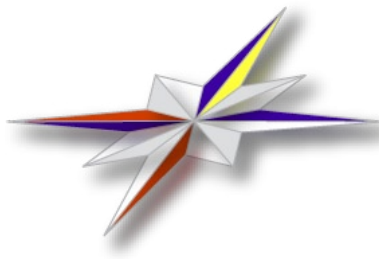
In order to design and optimize the ATLAS detector, to learn how to deal with unprecedented large amounts of data, and to prepare for the physics analysis, it is mandatory to simulate LHC data. ATLAS opted for a series of Data Challenges of increasing size and complexity to prepare and test the software in stages. The Data Challenges were executed by ATLAS research groups all over the world – Europe, USA, Canada, Japan, Taiwan, and Russia.

The Nordic ATLAS research groups were the first ones who could use entirely Grid tools to complete their share of the task. The testbed included 6 Linux clusters across Denmark (Copenhagen), Norway (Bergen, Oslo) and Sweden (Lund, Uppsala, Linköping). Despite having different operating systems and hardware characteristics, the clusters performed as a single virtual computing centre, having jobs distributed in an optimal way, and writing the output onto a dedicated storage area in Oslo. In total, 1300 jobs were successfully executed. The processing of 200 GB of input data produced 765 GB of output data, in 530 CPU-days. The average time to process a single event was 150 seconds. The ATLAS Data Challenge was successfully met, with the failure rate being negligibly small and no failures being attributed to the Nordugrid middleware. In the future, it is foreseen to continue running ATLAS Data Challenges on the Nordugrid, using upgraded middleware and increased computing resources.



Higgs particle decay simulation





Nordugrid future

The second phase of the Nordugrid project is foreseen to run in 2003-2004. The main task of the project will be to pursue research and development on a Nordic Grid solution, aiming at laying a foundation for a joint large-scale Nordic Grid Facility. Integration of the Nordugrid middleware with the middleware developed in other Grid projects will also be essential in order to facilitate Nordugrid becoming a seamless part of the emerging World Wide Grid.

While particle physics has been the main Grid customer in the first phase of the project, our purpose is to enlarge the user base and make the Grid tools available to all fields of science. This requires providing further information to, and co-operating with, scientists from other fields, and giving appropriate user support. To make this easier, we will continue improving the stability and documentation of the Nordugrid toolkit. In order to maintain an operational platform for users to test and execute their tasks, we will also continue operating the Nordugrid Testbed.



Contact us at contact@nordugrid.org if you wish to enquire how your computing application could be gridified!



www.nordugrid.org

