



NorGrid

THE NORWEGIAN GRID INITIATIVE

In the past decade, a number of software technologies have been developed to facilitate the automatic and transparent use of heterogeneous databases and commodity processor clusters that are connected through wide-area networks over large geographical distances, administrative domains and national borders. The infrastructures that are created this way are usually referred to as grids. Special software, so-called middleware, is required for the operation of the grid. Grids are arguably known as the next evolution in distributed computing and are used to optimize resource utilization and boost the aggregate capacity for data-intensive and loosely coupled parallel computations, sometimes by several orders of magnitude.

The objective of the NorGrid project is to provide grid-based services and to 'grid-enable' scientific applications on top of the infrastructure for computational science in Norway. The long term vision of the project is that many compute and storage resources will be accessed through a grid interface. The services and interfaces provide easy and secure access to distributed resources and help researchers create and participate in computational challenges of scope and size unreachable on single facilities alone.

SERVICES

Services that are important in a distributed infrastructure and where grid services contribute, include the following:

1. Resource management, brokering and scheduling.

Resource management enables the allocation of a job to a particular resource, monitor the status of the job while it is running and manage (e.g., cancel) it. Most grid systems include brokering and (meta-) scheduling software that decides, based on information about the available resources and a specification from the user, where and when a user task will be executed. Ultimately, the grid distributes work loads over the available resources, thereby reducing the turn-around time of jobs.

2. Distributed data management services. Intelligent sharing of data between users and applications is required in modern collaborative environments. Data sets need no longer be stored close to the user or compute resource (e.g., on a local file system). In a grid infrastructure, data for an application may be stored far away from where the application binaries reside. Grid-scheduling ensures that the data is at the appropriate location at the time that it is needed.

For the implementation of these services, virtualized environments are created that separate the underlying infrastructure from the applications that run on it. This separation allows applications and data to be dynamically placed and managed across a pool of resources. This way, the grid infrastructure can intelligently manage and shift

workload and data according to predefined policies. Vital components in the grid infrastructure are information services that maintain knowledge about resource availability, capacity, and current utilization.

ENABLING APPLICATIONS

An important task of NorGrid is to enable applications on the grid. This enabling must take into account data set sizes, data transfers over wide-area networks, replication of data, geographically distributed users and heterogeneous computing and storage resources. The most apparent advantage of grid environments today is for scientists that run large series of straightforward compute-intensive tasks with relatively simple input and output requirements. Applications that rely on access to distributed data collections (e.g., shared repositories) or meta-scheduling services can benefit as well.

WHO SHOULD USE A GRID ENVIRONMENT

- Researchers that have applications that can be partitioned in smaller tasks that can be distributed across multiple compute resources
- Researchers that have temporary peak-demands for compute or storage resources that cannot be handled by a single resource alone.
- Researchers with complex applications where user, applications, input/output data and/or compute resources are at different locations

A number of application areas already make use of grids. These include bio-informatics, earth sciences, chemistry, and high-energy physics. These applications are often based on compute-intensive processing of many independent data sets, the simulation of many independent scenarios, and Monte-Carlo simulation.

Researchers can apply for support to grid-enable applications and for access to compute and storage resources through www.norgrid.no.

INTERNATIONALLY

NorGrid aims to unite the operational grid efforts in Norway and reduce the management and organizational overhead of international cooperation around grids and related infrastructure projects. NorGrid enables Norway to participate in international grids and development work. Europe's largest grid is currently provided by EGEE. A key step in the development of future long-term sustainable grid infrastructures is the European Grid Initiative (EGI) that builds on National Grid Initiatives (NGIs). NGIs have recently been established in many countries. NorGrid is the Norwegian NGI.

Example: WLCG - the World-wide LHC Computing Grid

The European research centre CERN near Geneva hosts the Large Hadron Collider (LHC) that is built in a circular 27 km long tunnel that is designed to collide beams of accelerated protons or atomic nuclei and record the trajectories of elementary particles being created by these collisions. These collisions will produce roughly 15 Petabytes (15 million Gigabytes) of data annually, which thousands of scientists around the world will access and analyse. WLCG is the grid infrastructure to manage the data and distribute it from CERN through a layered network to eleven (Tier-1) centres world-wide that store and process the data and make it available to more than one hundred other (Tier-2) centres where the data is further analyzed. The Nordic Data Grid Facility (NDGF) is a Tier-1 centre and builds on the resources provided by the Nordic countries, including NorGrid.

NorGrid partners

The NorGrid consortium currently exists of the Norwegian University of Science and Technology (NTNU), the University of Bergen (UiB), the University of Oslo (UiO), the University of Tromsø (UiT), UNINETT and UNINETT Sigma (coordinator).

Links

NorGrid: www.norgrid.no
Notur: www.notur.no
EGEE: www.eu-egee.org
EGI: www.eu-egi.org
NDGF: www.ndgf.org
WLCG: lcg.cern.ch/LCG