

Grid Computing on Nordugrid

Josva Kleist

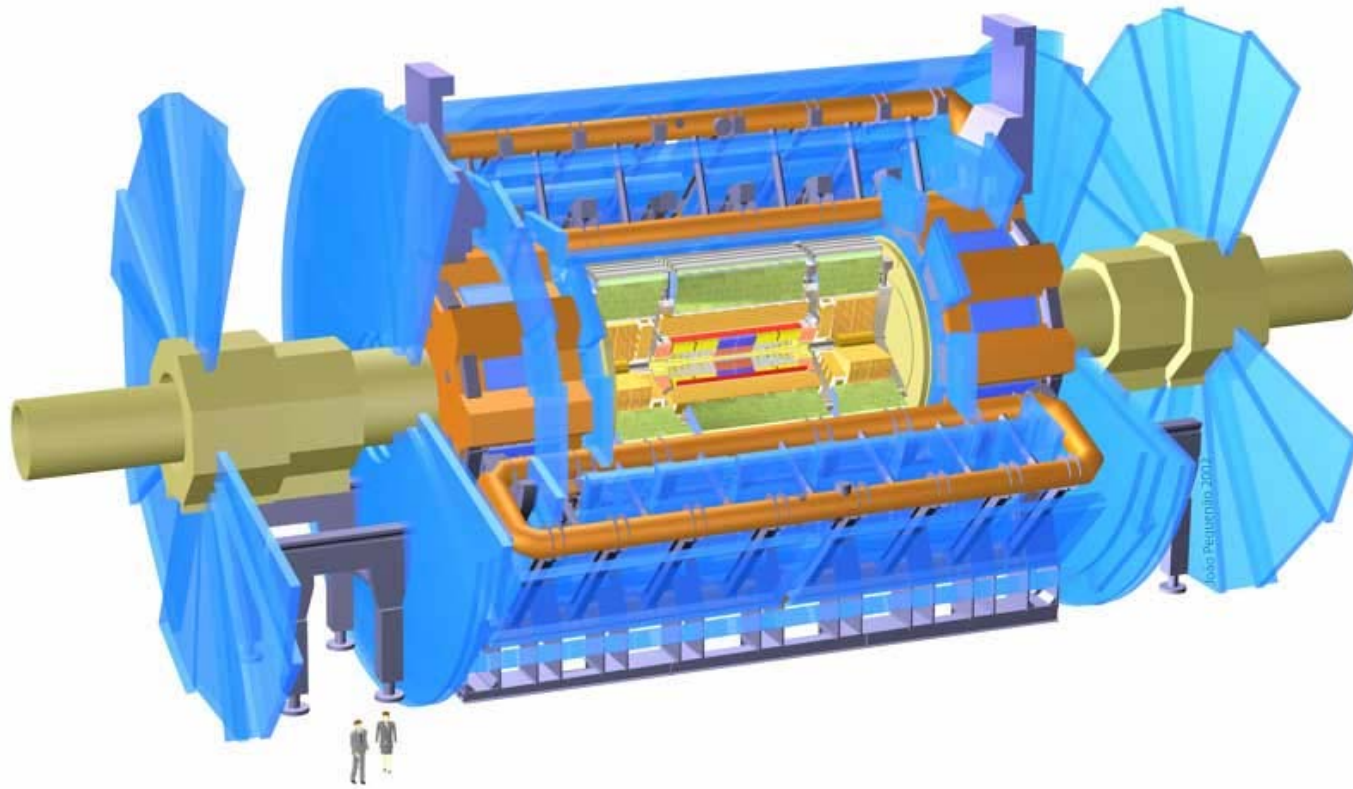
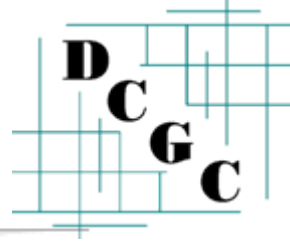
Danish Center for Grid Computing

(www.dcgc.dk)

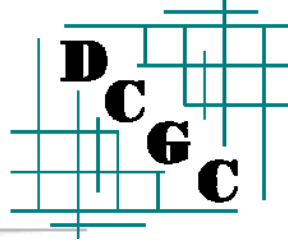
&

Aalborg University

The Atlas experiment

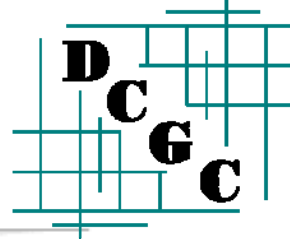


Agenda

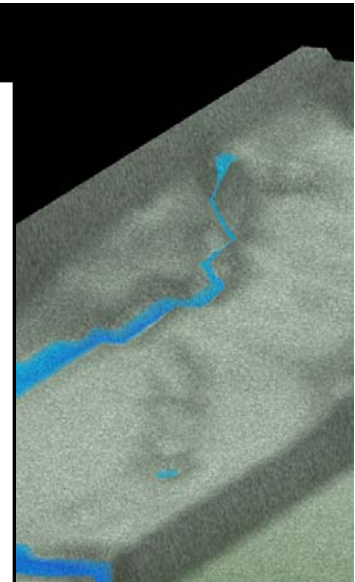
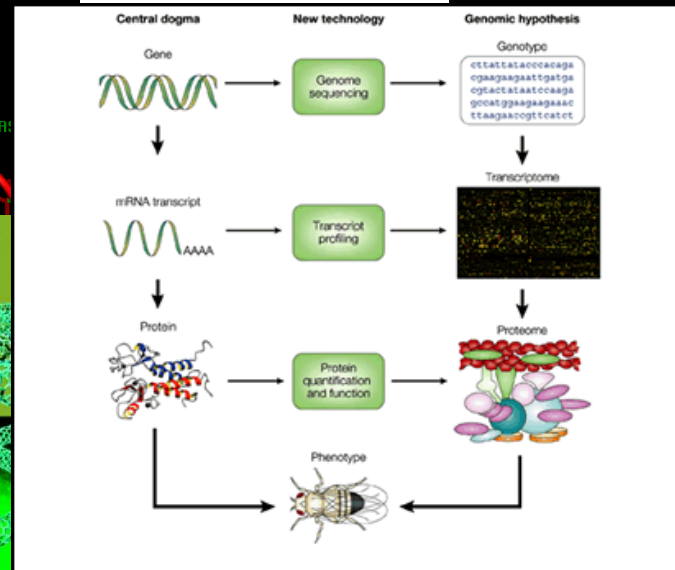
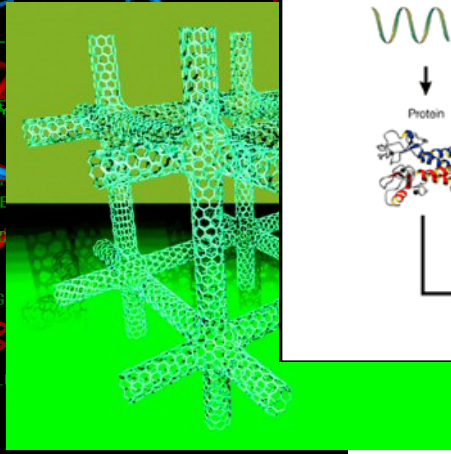
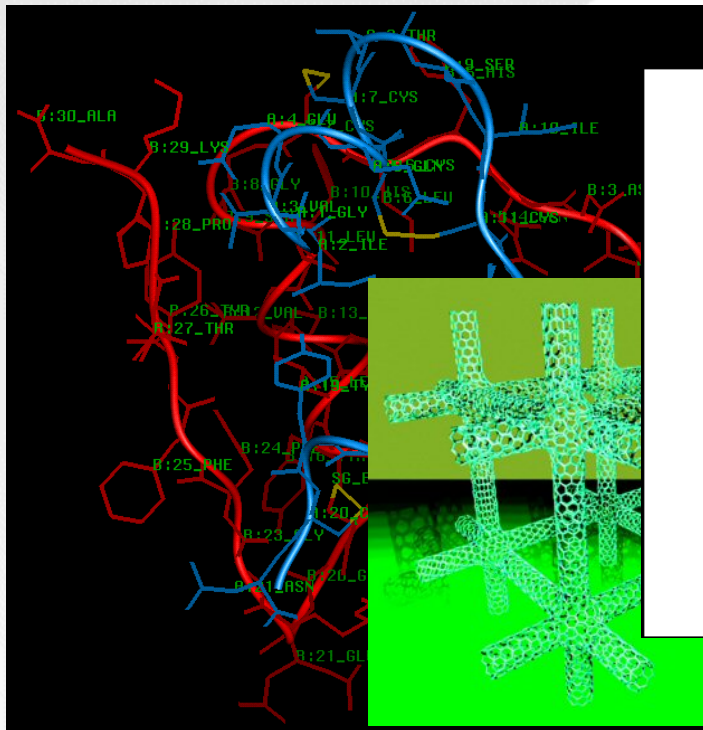


- Grid Computing defined
- NorduGrid ARC – overall structure and use
- Demo
- Distributed model checking on NorduGrid
- Future

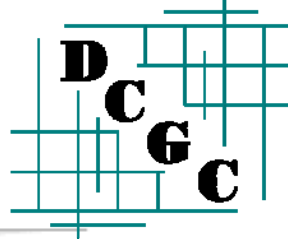
E-science



”Science (increasingly) done through distributed global collaborations enabled by the Internet, using very large data collections, tera-scale computing resources and high performance visualisation.”

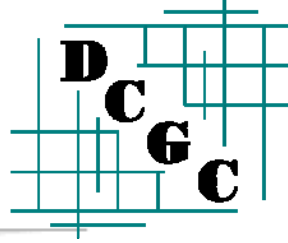


The grand vision



A huge virtual distributed computer.

Definition



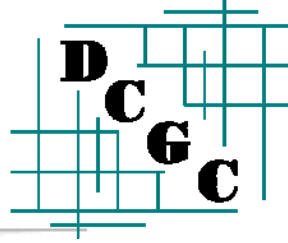
“A computational grid is a hardware and software infrastructure that provides dependable, consistent, pervasive, and inexpensive access to high-end computational capabilities.”

The Grid – a blueprint for a new computing infrastructure, 1998

DMC to the *people* need a Grid because

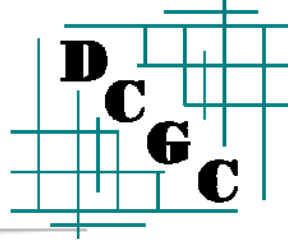
- Ordinary users of MC don't have the needed HW
- The software is too difficult to install and maintain

Challenges



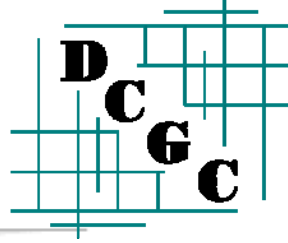
- Make hardware owned by *different* organizations available to *non-members* of that organization.
- In such a way that normal operation of the equipment can continue.
- In such a way that the organization still can control who gets access.
- In such a way that we can control who gets access to specific pieces of data.
- In such a way that operations can be performed anonymously.
- And still charge for the use of hard- and software.

Challenges



- Resource allocation and scheduling
- Authentication and authorization
- Protection
- Control
- Accounting

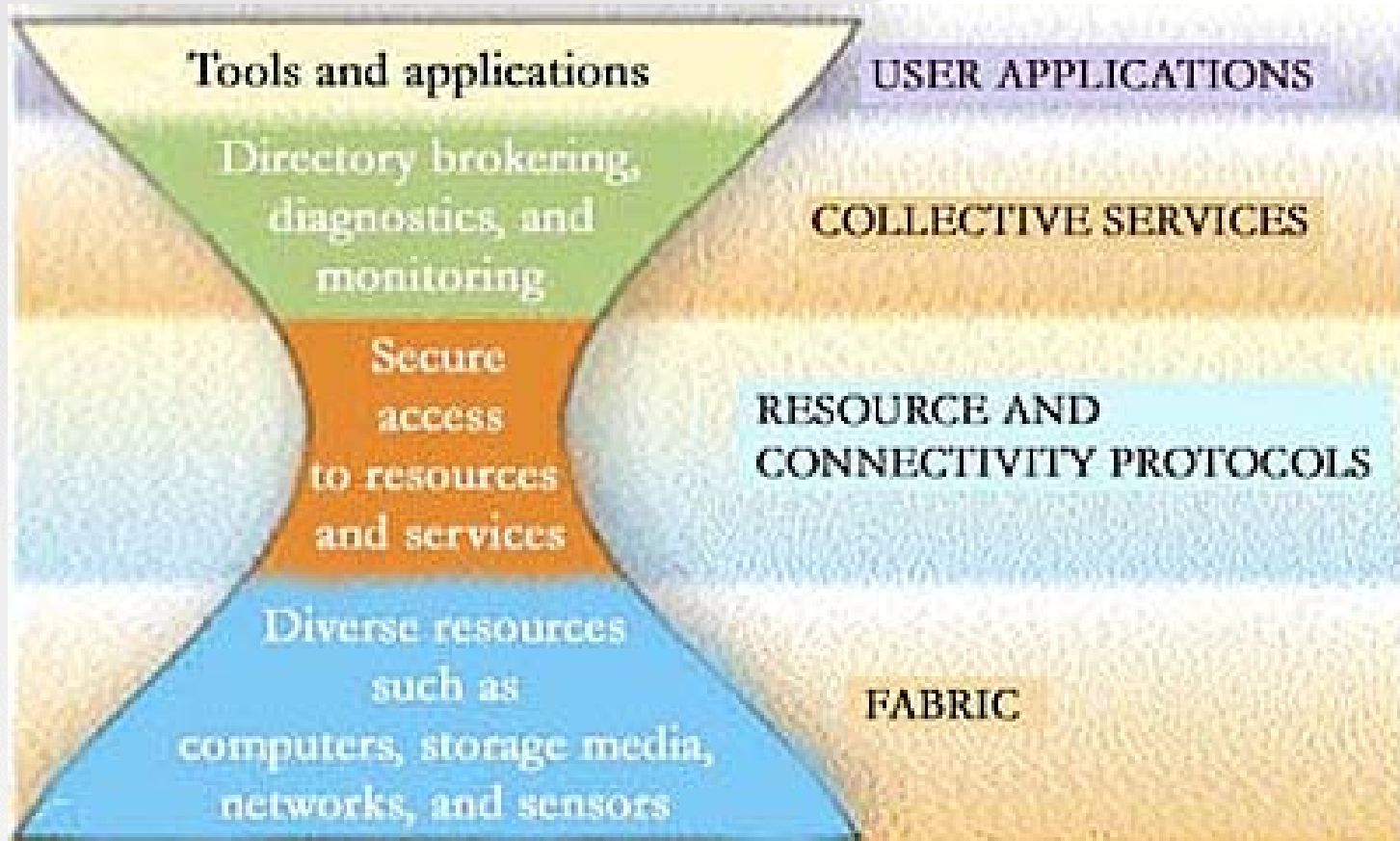
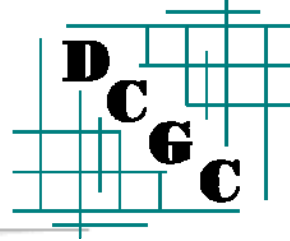
Globus



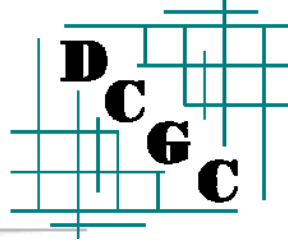
- An open source software toolkit used for building grids.
- Includes software services and libraries for resource monitoring, discovery, and management, plus security and file management.

Web: www.globus.org

The globus model



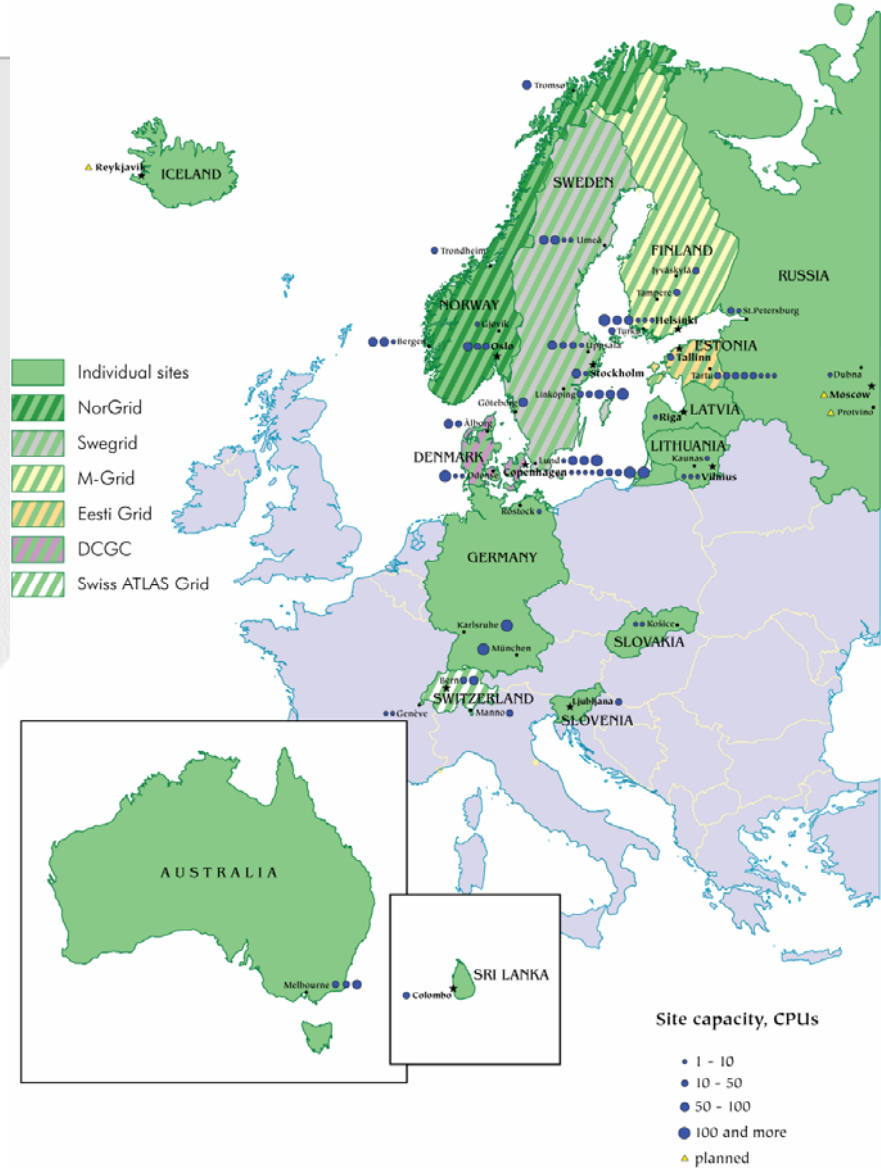
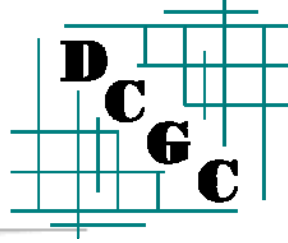
NorduGrid



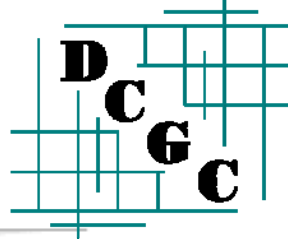
- NorduGrid is a collaboration between a number of universities mostly located in the Nordic countries.
- NorduGrid Advanced Resource Connector is:
 - A Globus-based Grid middleware solution
- NorduGrid is a production Grid
 - Approximately 5000 CPUs
 - Approximately 75 TB of storage

Web: www.nordugrid.org

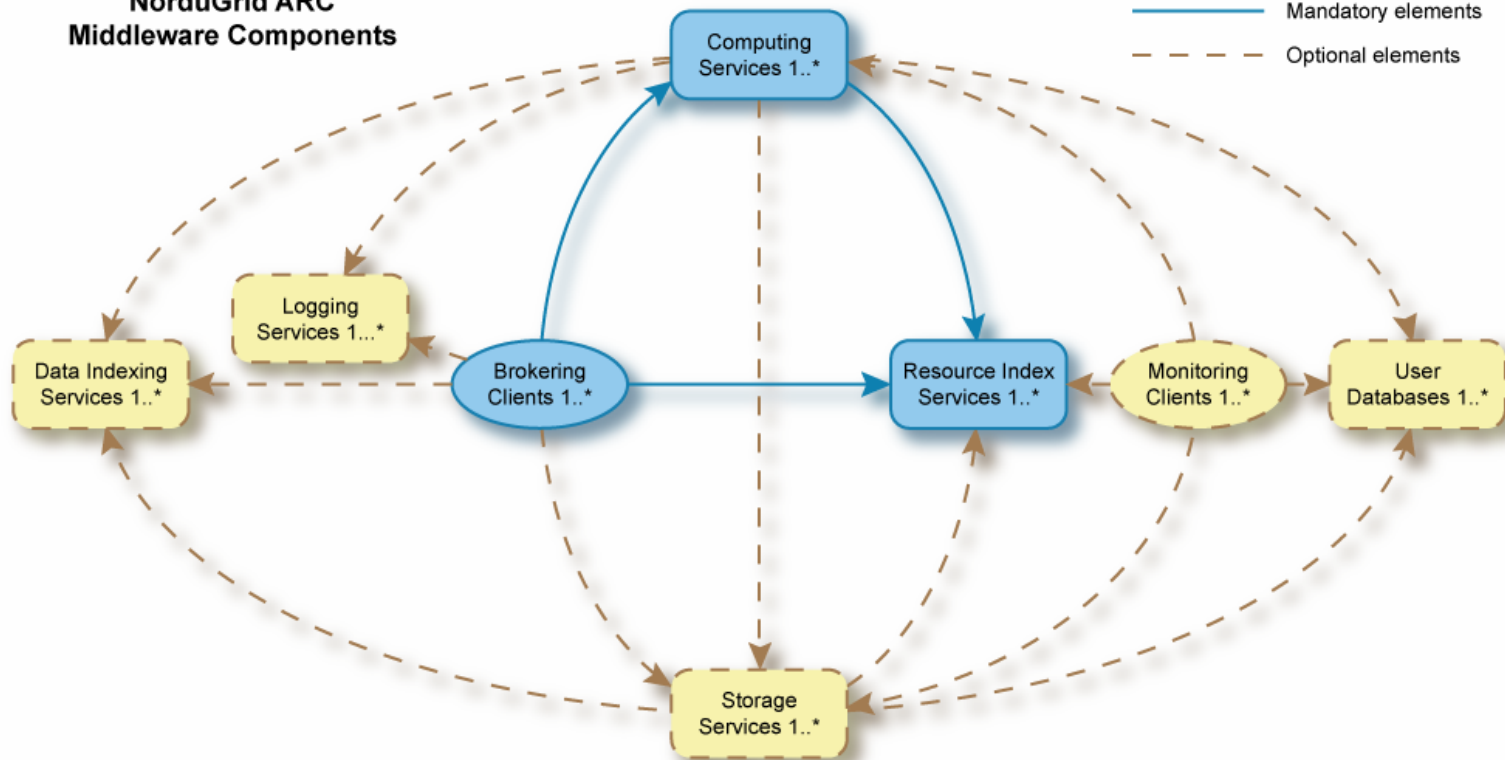
The Nordic Grid



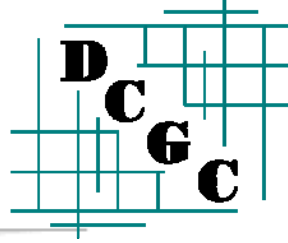
Demo



NorduGrid ARC Middleware Components

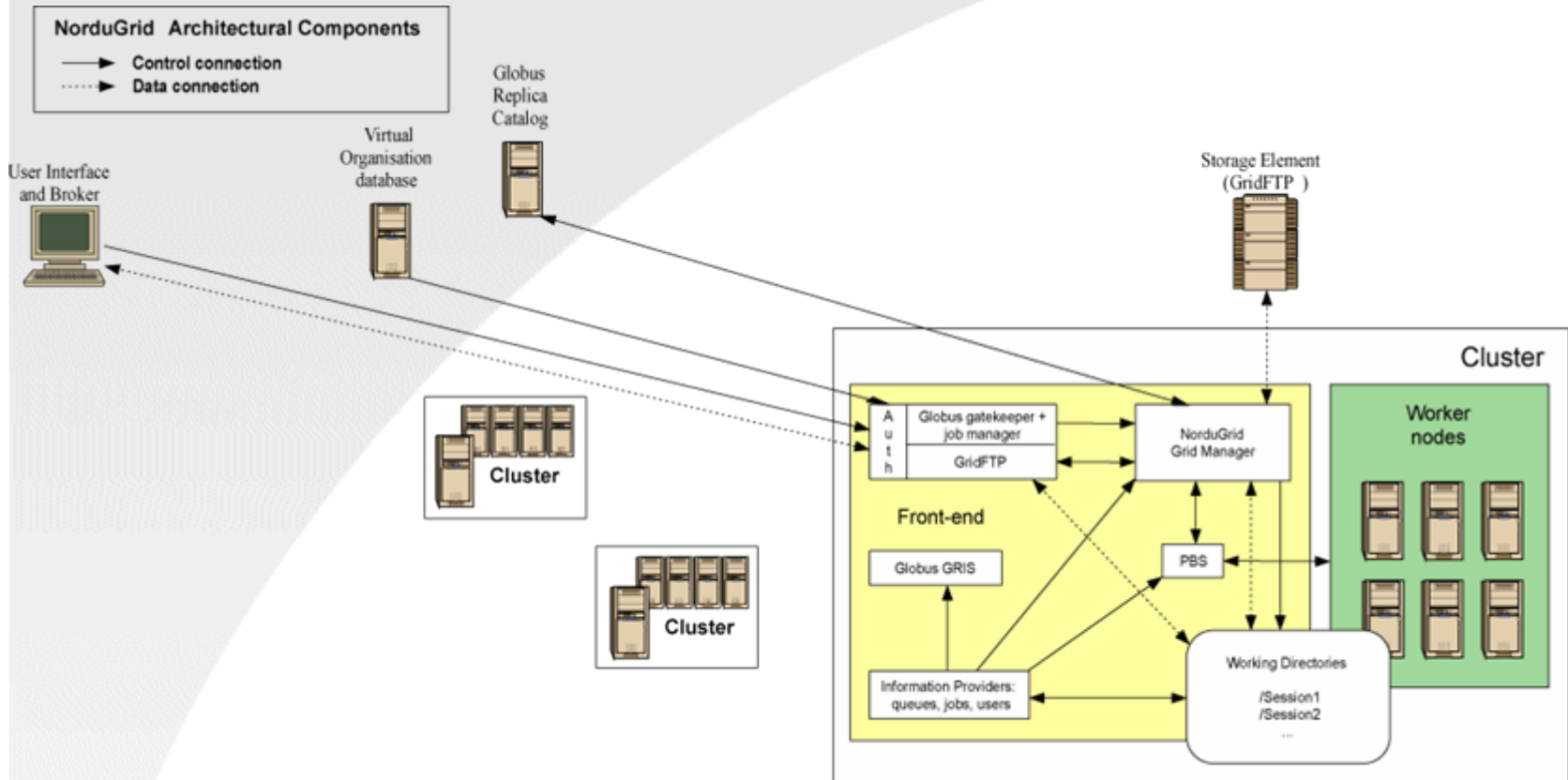
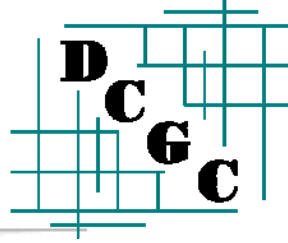


ARC Assumptions



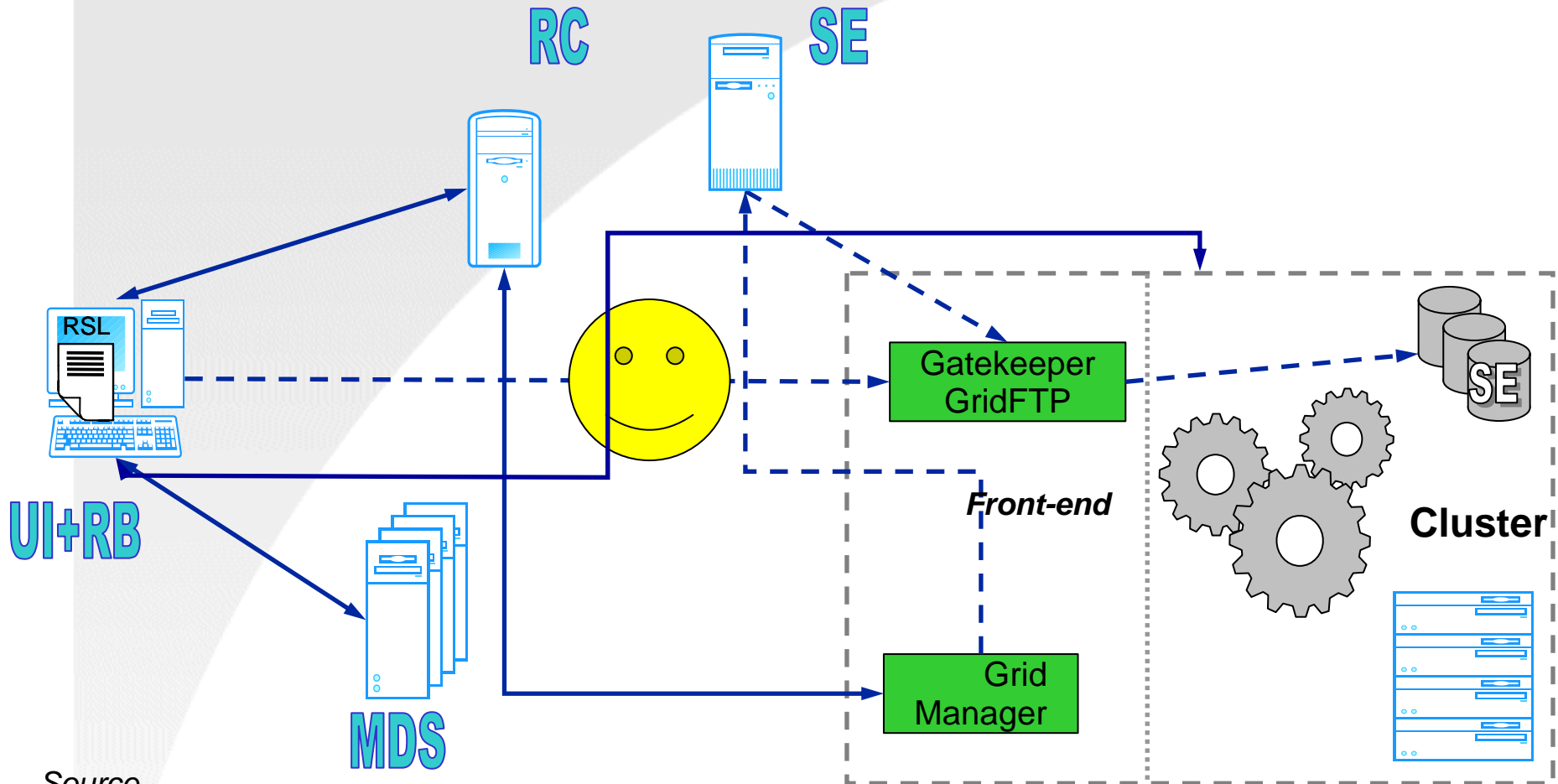
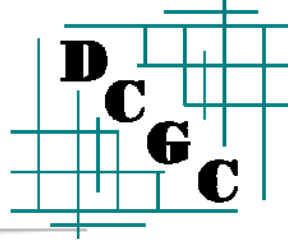
- It should only be necessary to install Grid software on a single *front-end* machine.
- No/few modifications to the existing hardware/software configuration should be necessary.
- Bottlenecks known to exist in other Grid solutions should be avoided.

ARC Components



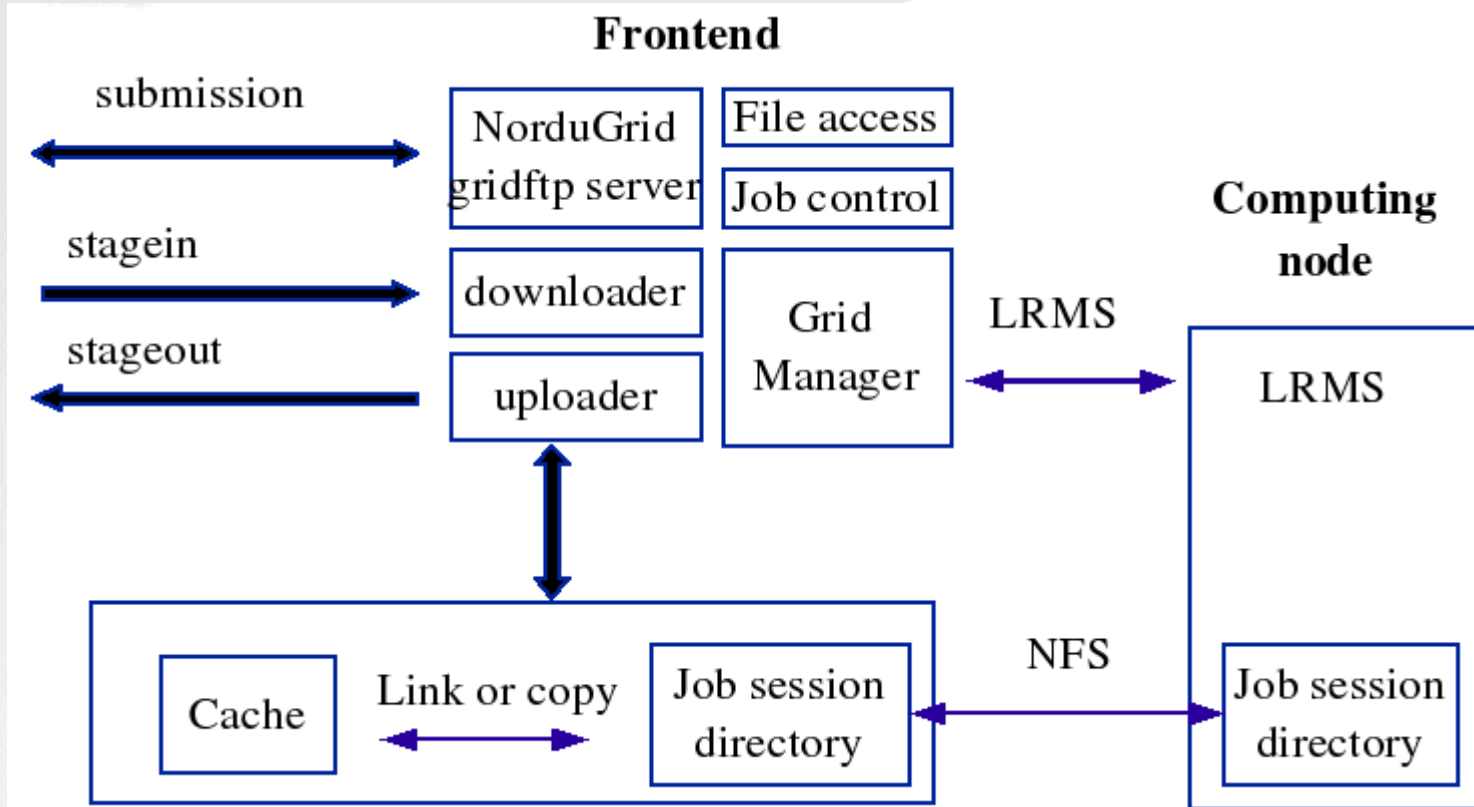
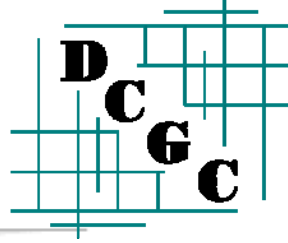
October 2002

Workflow

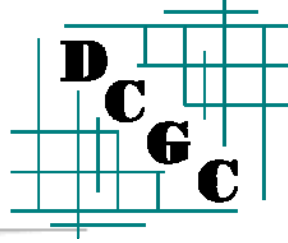


Source
NorduGrid.org

Front-end

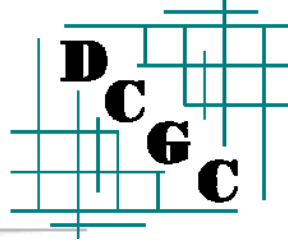


The user-interface



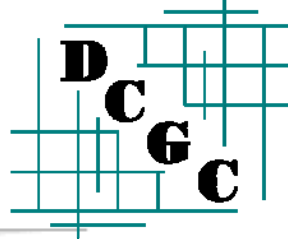
<i>ngsub</i>	to submit a task
<i>ngstat</i>	to obtain the status of jobs and clusters
<i>ngcat</i>	to display the stdout or stderr of a running job
<i>ngget</i>	to retrieve the result from a finished job
<i>ngkill</i>	to cancel a job request
<i>ngclean</i>	to delete a job from a remote cluster
<i>ngrenew</i>	to renew user's proxy
<i>ngsync</i>	to synchronize the local job info with the MDS
<i>ngcopy</i>	to transfer files to, from and between clusters
<i>ngremove</i>	to remove files

Broker



- The user must be authorized to use the cluster and the queue
- The cluster's and queue's characteristics must match the requirements specified in the xRSL string (max CPU time, required free disk space, installed software etc)
- If the job requires a file that is registered in a Replica Catalog, the brokering gives priority to clusters where a copy of the file is already present
- From all queues that fulfill the criteria one is chosen randomly, with a weight proportional to the number of free CPUs available for the user in each queue
- If there are no available CPUs in any of the queues, the job is submitted to the queue with the lowest number of queued job per processor

DMC on NorduGrid



- Distributed Uppaal is available on NorduGrid, through a special *runtime environment*.
- No integrations with the Uppaal UI, but users specify queries in their Grid *xRSL* specification

A real DMC Grid should have

- Grid access integrated into the single processor UI
- The Grid should support interactive jobs
- Support advance reservations and co-scheduling