

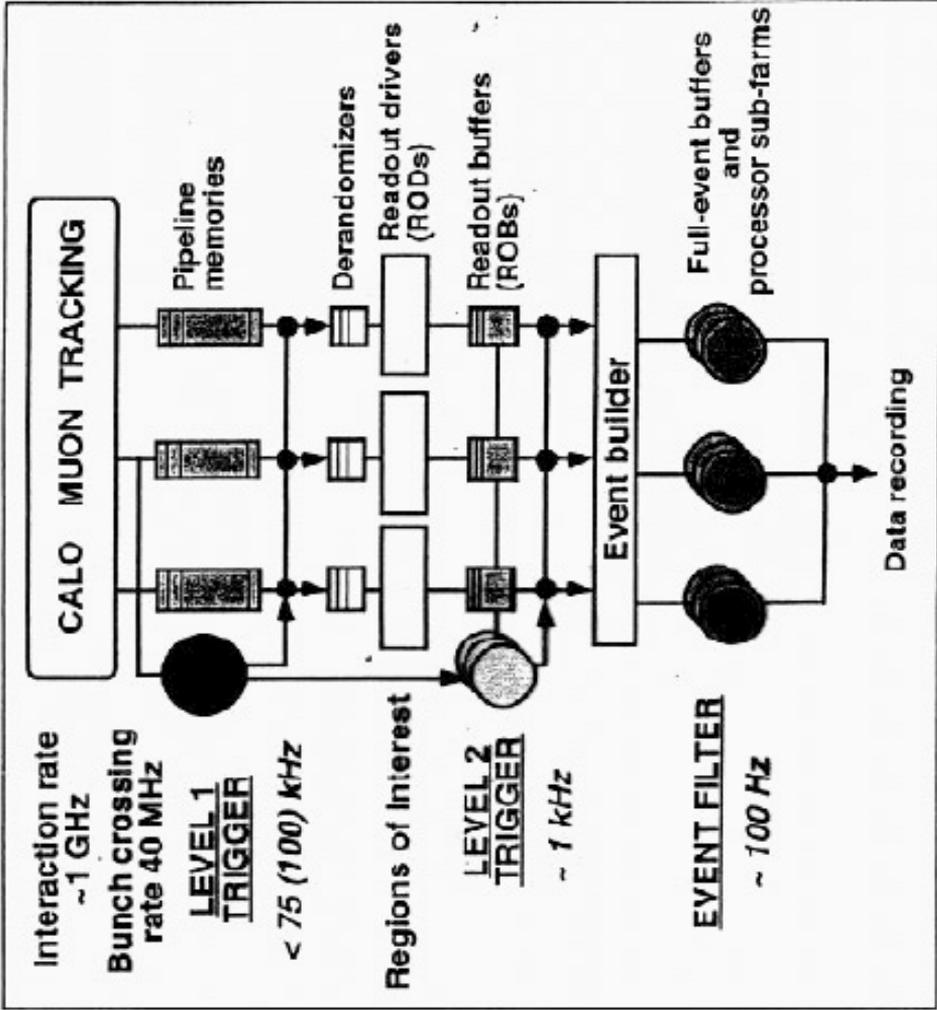
NorduGRID

*A Nordic test bed for wide area
computing and data handling*

Presented by John Renner Hansen
The Niels Bohr Institute

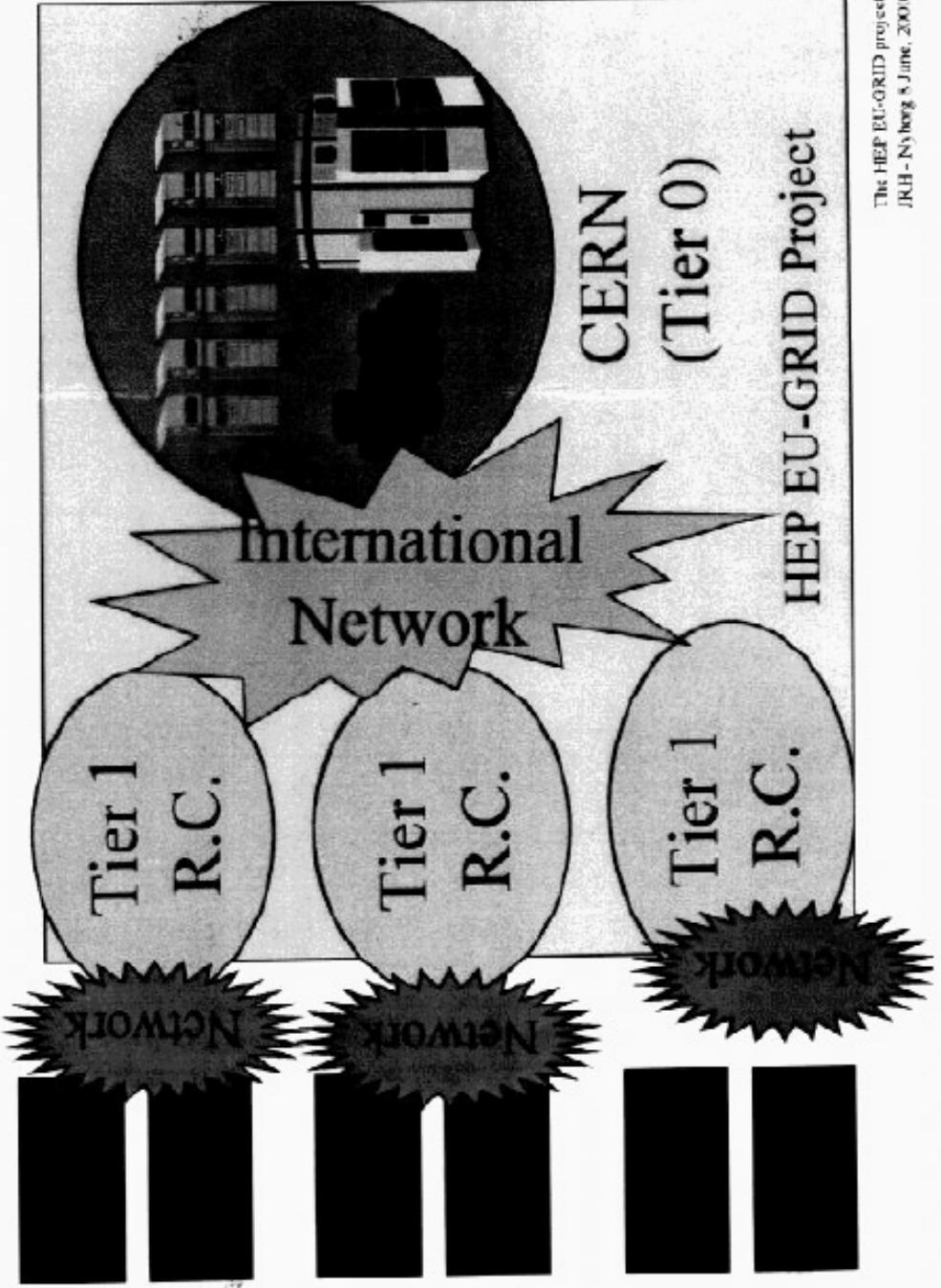
23 February 2001

NOS-N Grid seminar





LHC Regional Centers

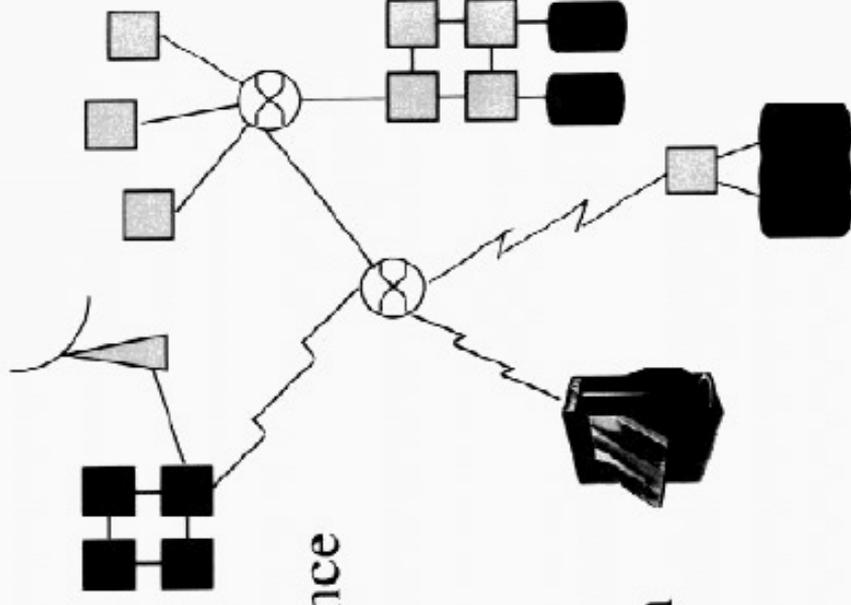




The Grid

“Dependable, consistent, pervasive access to [high-end] resources”

- Dependable: Can provide performance and functionality guarantees
- Consistent: Uniform interfaces to a wide variety of resources
- Pervasive: Ability to “plug in” from anywhere



GRID

Edited by Ian Foster
and Carl Kesselman





Atlas Estimate & Rules of Thumb

- **Tier 1 Center in '05 should include ...**
 - 30,000 SPECint95 for Analysis
 - 10-20,000 SPECint95 for Simulation
 - 50-100 TBytes/year of On-line (Disk) Storage
 - 200 TBytes/year of Near-line (Robotic Tape) Storage
 - 100 MB^{bytes}/sec connectivity to CERN
- **Assume no major raw data processing or handling outside of CERN**



Components of Model: Tier 1

- **Full Function Facility**
 - **Dedicated Connectivity to CERN**
 - **Primary Site for Storage/Serving**
 - **Cache/Replicate CERN data needed by outside institutes**
 - **Archive and Serve WAN all data of interest to outside institutes**
 - **Computation**
 - **Primary Site for Re-reconstruction (perhaps only site)**
 - **Major Site for Simulation & Analysis (~2 x Tier 2)**
 - **Repository of Technical Expertise and Support**
 - **Hardware, OS's, utilities, and other standard elements of ATLAS**
 - **Network, AFS, GRID, & other infrastructure elements of WAN model**

Application to the Nordunet-2 Programme

A Nordic test bed for wide area computing and data handling

27 October 2000

Project co-ordinator: Associate Professor John Renner Hansen email: renner@nbi.dk
Niels Bohr Institutet for Astronomi, Fysik og Geofysik
Blegdamsvej 17
Dk-2100 Copenhagen Ø
Ext: +45 35 32 53 27

Participating institutes:

Denmark:

Research Center COM at the Danish Technical University

Department of Computer Science at the University of Copenhagen (DIKU)

Niels Bohr Institutet for Astronomi, Fysik og Geofysik (NBIIAFG) at the Univ. of Copenhagen

Finland:

University of Helsinki

Norway:

Department of Physics at the University of Bergen and

Department of Physics at the University of Oslo

Sweden:

Department of Physics at the University of Lund

Department of Physics at the University of Uppsala

Department of Physics at the University of Stockholm

Department of Physics at KTH

Parallel Data Centrum in Stockholm

NorduGRID objectives.

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 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 petabyte storage requirements and processing power in the order of multi tera-flop.

Work package composition

The following six work packages are foreseen to cover the work:

WP1: Test bed hardware:

The main purpose of this work package is to insure smooth hardware integration by

- understanding existing network infrastructure in the Nordic countries
- propose improved network resources where necessary
- understanding existing farm hardware in order to provide a heterogeneous GRID
- providing help for the installation of new farm hardware whenever necessary for the progress of the project

Sites involved in the test bed are foreseen to join gradually, with links to be established in approximately the following order:

- | | |
|---|----------------|
| • Lund - NBI | January 2001 |
| • Lund - Stockholm ² | February 2001 |
| • NBI - Lund-Stockholm | March 2001 |
| • Uppsala - Stockholm | April 2001 |
| • Bergen/Oslo - NBI | June 2001 |
| • Bergen/Oslo - NBI - Lund - Stockholm - Uppsala | August 2001 |
| • Bergen/Oslo - NBI - Lund - Stockholm - Uppsala - Helsinki | September 2001 |

WP2: Farm software:

Provide system and management tools for farms:

- Installation of OS (Unix dialect) and applications and utilities, i.e. Globus toolkit, SSLcay, OpenLDAP, Condor, JDK, Monarch simulation tools etc.
- Configuration; RSL, scheduling
- Monitoring;
- Alarms
- Resource discovery

Provide access from the GRID

- Support external requests
- Publication of available resources
- Security: Authentication / Authorization
- Policies: Allocation / Priorities / Estimation / Cost
- Scheduling
- Reservation
- Guarantees Resource availability

² Stockholm means one or more of the participating institutes in the Stockholm area, i.e. FDC, KTH and Stockholms Universitet.

WP3: Grid Monitoring Services

Specify, develop, integrate and test tools and infrastructure to enable end-user and administrator access to status and error information in a Grid environment.

Goals are to permit both job performance optimisation as well as allowing for problem tracing, crucial to facilitating high performance Grid computing, taking into account issues as

- Scalability - large numbers of resources and users
- Architectural - optimisations
- Agent Technology Information Discovery - via directory services
- Security (for info access, etc.)

WP4: Grid Data Management

Specify, develop, integrate and test tools and middle-ware infrastructure to investigate coherent data management and shared Pbyte-scale information volumes in high-throughput production-quality grid environments. The work package will integrate general-purpose information sharing solution with scalability, uniformity, transparency and heterogeneity.

It shall allow to securely access massive amounts of data in a universal global name space, to move and replicate data at high speed from one geographical site to another, and to manage synchronisation of remote replicas.

Provide system and management tools for data handling:

- generic interfacing to heterogeneous mass storage
- for automated wide-area data caching distribution will act according to dynamic usage software patterns

WP5: User applications and benchmarks

Define and conduct exhaustive tests of the Nordic GRID environment. Scalability, reliability and usability should be demonstrated.

- Definition of benchmark tests
- Preparation of the test bed interfacing with the HEP experimental off-line environment, evaluation of results and optimisation of the prototype parameters
- Development of the experiment off-line software to take advantage of a distributed heterogeneous computing environment.

Examples of applications which can be executed over the network; event generation, Geant4 simulation or test-beam data analysis.

WP6: Dissemination

Preparing and maintaining web sites with project related information for private and public use.

- Prepare two workshops, 2001-Q1 and 2002-Q2
- Organize video conferences on the bi-monthly basis to improve communication between sites

Work package assignment:

WORK PACKAGE	INSTITUTES INVOLVED	MAIN RESPONSIBLE PERSON
1. Test bed hardware	NBI Lund SU	Post Doc 1
2. Farm software	NBI Lund KTH SU Finland COM DIKU	Post Doc 2
3. Grid Monitoring Services	Lund COM DIKU	Post Doc 2
4. Grid Data Management	COM DIKU	Post Doc 2
5. User applications and benchmarks	NBI Lund KTH SU Finland Bergen Oslo	Post Doc 1
6. Dissemination	All institutes	Post Doc 1

Project plan and deliverables:

The project is assumed to start January 1st 2001 and last for 27 month.

Deliverables:

January 1st 2001

A password protected Web page for internal communication.

The Web page contains minutes, status reports, lists of scheduled meetings and video conferences, internal papers, etc.

Work package(s) involved: 6

February 1st 2001

A public web page for general project information to a wider community.

The page contains project status information, information on how to become a test user and links to relevant sites, i.e. GLOBUS, DATAGRID, GRIPHYN, etc.

Work package(s) involved: 6

February 2001

Globus workshop in Lund or Copenhagen

Work package(s) involved: 2 and 5

March 1st 2001

Work document on test programme and bench marking.

To be prepared at the workshop in Lund or Copenhagen February 2001.

Work package(s) involved: 5

May 1st 2001

A test bed connecting Lund and Copenhagen based on the GLOBUS software package.

HEP simulation production can be submitted to both farms from the two sites.

Users from other fields can get access for limited periods and with access to limited resources.

Work package(s) involved: 1, 2, 3 and 5

September 1st 2001

A test bed extended to farms from all participating countries

Work package(s) involved: 1, 2 and 3

October 1st 2001

Job submission is executed from the protected Web pages

Performance monitoring is available on the public Web site.

Work package(s) involved: 2

January 1st 2002

Access control and full security established.

Work package(s) involved: 2 and 3

March 1st 2002

Global accounting system

Work package(s) involved: 2 and 3

June 1st 2002

A test bed includes all sites with full functionality, according to availability of machines and network connections.

Multi disciplinary access is open.

Work package(s) involved: 2, 3, 5 and 6

December 1st 2002

Processing follows data

Users do not have to know where data actually are stored. Data are automatically moved or mirrored to the place where they most often are used.

Work package(s) involved: 3, 4

July 1st 2003

Final report.

Workshop in Bergen, Helsinki, Oslo, Stockholm or Uppsala in February 2003, to prepare work for final document.

Work package(s) involved: All

Nordic perspectives

Nordic countries, as member states of CERN, actively participate in the LHC project as a whole, and in the DATAGRID in particular. Universities in Denmark, Finland, Norway and Sweden are getting involved in the development of the software tools and applications. A very important step towards deployment of the High Energy Physics distributed computing network would be creation of a Nordic chain, which initially will connect universities participating in the LHC project and expand further to be used by other research groups and applications. This chain shall serve as a test bed for the DATAGRID project.

Inter Nordic Network connections and beyond.

The main hardware building blocks are local clusters of PC's and international high bandwidth data connections between the clusters. Most of the participating institutes will contribute between 2 and 40 PC's to the GRID. Fast and stable network connection between the sites is essential for the success of the project. Three such international connections are required; between Denmark and Sweden, between Denmark and Norway and between Sweden and Finland. These could be implemented on the new NORDUnet 2.5 Gbit research network, presently being commissioned. Requirements and conditions should be discussed with the NORDUnet management in the near future. In an initial phase one could make use of the now inactive 155 Mbit radio chain between Copenhagen and Lund. Conditions for reopening and use of this facility are to be discussed with the owner.

A project between NBI and CERN aims to establish a fast permanent link between Geneva and Copenhagen for exclusive use by the two partners. A Nordic network will be a natural extension to this link.

Project implementation.

Implementation of the Nordic DATAGRID test bed is foreseen to proceed within 6 work packages. The task distribution between packages overlaps consciously with the work packages in the CERN DATAGRID proposal, giving optimal utilisation of resources, exchange of experience and optimisation for participation in the next EU application from CERN. A new call for proposals to EU is expected in 2001.

Two workshops in February 2001 and February 2002 are foreseen to establish the detailed work programme for the project and the final set of benchmark tests respectively.

Results will be made available for presentation at the annual NORDUnet meeting, at "Computing in High Energy Physics 2001" and at similar meetings and conferences.

Financial request from Nordunet-2

Of the six work packages WP2, WP3 and WP4 are the most demanding in terms of understanding new software methodologies, coordination of many institutes and with the other work packages. A wide area maintenance programme must be created. This will require full time participation of at least two persons to insure that these packages deliver as foreseen. We therefore request resources for two persons at post doc. level. The positions will be announced for open competition in the Nordic countries. For geographical reasons, one will be based in Lund or Copenhagen, the other in Stockholm or Uppsala. Apart from being responsible for one or two work packages it is the aim that Post Doc. 1 (PD1) provides help to Lund plus the Danish and Norwegian groups, PD2 shall provide help to the remaining Swedish and the Finnish groups. With extensive use of network communication tools and videoconference, most maintenance can be performed remotely.

2 post doc. in 2 years	800 kkr/year	1.6 Mdkr
------------------------	--------------	----------

A moderate amount of travel for the two Post Doc. must be foreseen, in order to keep contact with the other groups and to participate in CERN DATAGRID meetings and conferences, 2 x 30 kdkr/year. We would also like to have the possibility to invite speakers to the two workshops, 20 kdkr/year, as well as supporting student participation, 20 kdkr/year.

Travel	100 kkr/year	0.2 Mdkr
--------	--------------	----------

A few hardware component and software licences as well as some payment for high-speed international data connection is foreseen.

Hardware and software licences	100 kkr/year	0.2 Mdkr
International data connection	100 kkr/year	0.2 Mdkr

Summary request from Nordunet-2

2 post doc. in 2 years	800 kkr/year	1.6 Mdkr
Running cost		0.6 Mdkr
Travel	100 kkr/year	
Hardware	100 kkr/year	
International data connection	100 kkr/year	
Total over two years		2.2 Mdkr