"We will probably see the spread of ‘computer utilities’, which, like present electric and telephone utilities, will service individual homes and offices across the country.”
Len Kleinrock (1969)

“A computational grid is a hardware and software infrastructure that provides dependable, consistent, pervasive, and inexpensive access to high-end computational capabilities.” I. Foster, C. Kesselmann (1998)

"The sharing that we are concerned with is not primarily file exchange but rather direct access to computers, software, data, and other resources, as is required by a range of collaborative problem-solving and resource-brokering strategies emerging in industry, science, and engineering. The sharing is, necessarily, highly controlled, with resources providers and consumers defining clearly and carefully just what is shared, who is allowed to share, and the conditions under which sharing occurs. A set of individuals and/or institutions defined by such sharing rules what we call a virtual organization.” I. Foster, C. Kesselmann, S. Tuecke (2000)
Introduction: Grid Literature

Books:
• I. Foster, C. Kesselmann: *The Grid: Blueprint for a New Computing Infrastructure*, Morgan Kaufmann Publisher Inc. (1999)

Articles:

Source:

Introduction: Talks@DESY


V. White (FNAL): *Computing at Fermilab: from Run II to LHC and the Grid* (7 April 2003).


Introduction: This Talk

- Give an(other) introduction into the Grid
- Fill the gap between high level information and actual installation
- Describe how DESY is involve
- Describe and discuss the Grid Testbed
- Provide a look-and-feel of its functionality
- Give a perspective for the future

Introduction: Contents

- Introduction
- Grids
- Grid Activities at DESY
  - Grid Testbed
  - dCache
  - ILDG
- EGEE
- Conclusions
Grids?

- Compute Grids
- Data Grids
- Science Grids
- Access Grids
- Knowledge Grids
- Bio Grids
- Sensor Grids
- Cluster Grids
- Campus Grids
- Tera Grids
- Commodity Grids
- Funding Concept?
- Marketing Slogan?

Grids: What is a Grid?


“A Grid is a system that:

coordinates resources which are not subject to centralized control …

integration and coordination of resources and users of different domains vs. local management systems (batch systems)

… using standard, open, general-purpose protocols and interfaces …

standard and open multi-purpose protocols vs. application specific system

… to deliver nontrivial qualities of services.”

coordinated use of resources vs. uncoordinated approach (web)
Grids: Types

Data Grids:
- provide transparent access to data which can be physically distributed within Virtual Organizations (VO)

Computational Grids:
- allow for large-scale compute resource sharing within Virtual Organizations (VO)

Information Grids:
- provide information and data exchange, using well defined standards and web services

Grids: Schematics
Grids: HEP Motivation

Data Grid Hierarchy (CMS)

- **Tier 1**: Online System
  - Bunch crossing per 15 ns each
  - 10^5 triggers per second
  - Event size ~ 1 MB/second

- **Tier 2**: Tier 0
  - CESN Computer Center
  - 1 TB/s
  - 100 Mbps/sec

- **Tier 3**: Tier 2
  - Europe Regional Center
  - 100 - 1000 Mbps/sec
  - 1.2 Gbps/sec
  - 20 Tbps/sec

- **Tier 4**: Tier 3
  - Workstations
  - Physicists work on analysis "channels".
  - Each institute has ~10 physicists working on one or more channels.
  - Data for these channels should be cached by the institute server.

Grid Projects

- **EGEE**: Enabling Grids for Science in Europe
- **ILDG**: International Long-Distance Grid
- **DataGRID**: Data Grid
- **LCG**: Large Circular Grid
- **IVERG**: Integrated Virtual Environment for Research Grids
HEP Grids Worldwide

HEP Grids in Europe
**Grid: Middleware**

Globus: ([http://www.globus.org](http://www.globus.org))
- Toolkit
- Argonne, U Chicago

- Project to develop Grid middleware
- Uses parts of Globus
- Funded for 3 years (1.4.2001 - 31.3.2004)

LCG (LHC Computing Grid): ([http://cern.ch/lcg](http://cern.ch/lcg))
- Grid infrastructure for LHC production
- Based on stable EDG versions; other approaches in the US (VDT)

- Grid infrastructure for enhanced science in Europe
- 4 years, first phase funded (1.4.2004 – 31.3.2006)

---

**Certificates**

- Authorization and authentication are essential ingredients of Grids
- A certificate is an encrypted electronic document, digitally signed by a Certification Authority (CA)
- A Certificate Revocation List (CRL) published by the CA
- For Germany GridKa issues certificates on request (needs ID copy)
- Contacts at DESY: R. Mankel, A. Gellrich

- Users, hosts, and services must be certified
- The Globus Security Infrastructure (GSI) is part of the Globus Toolkit
- GSI is based on the openSSL Public Key Infrastructure (PKI)
- It uses X.509 certificates
- The certificate is used via a token-like proxy
- In my case: /O=GermanGrid/OU=DESY/CN=Andreas Gellrich
A Virtual Organization (VO) is a dynamic collection of individuals, institutions, and resources which is defined by certain sharing rules.

- Technically, a user is represented by his/her certificate
- The collection of authorized users is defined on every machine in `/etc/grid-security/grid-mapfile`
- This file is regularly updated from a central server
- The server holds a list of all users belonging to a collection

It is this collection we call a VO.

- The VO a user belongs to is NOT part of the certificate!
- A VO is defined in a central list, e.g. a LDAP tree
- In our case there is a VO=DESY on a central server

DESY does not participate in LHC and is therefore not in LCG

- Germany’s main hub (regional centre: Tier-1) in LCG is GridKa at FZ Karlsruhe
- DESY’s technical Grid activities:
  - Grid Testbed
  - dCache
  - International Lattice Data Grid (ILDG)
- DESY’s (official) way to the Grid:
  - Enabling Grids for E-Science in Europe (EGEE)
  - D-GRID Initiative
**Grid Testbed: Project**

- The Grid Testbed was initiated by IT and H1 (R. Gerhards, A. Gellrich)
- It is an open initiative to study Grids ([mailto:grid-dev@desy.de](mailto:grid-dev@desy.de))
- Main developers: M. Vorobiev (H1), J. Nowak (H1), D. Kant (QMUL), A. Gellrich (IT)
- Further members: A. Campbell (H1), B. Lewendel (HERA-B), C. Wissing (H1 Do), S. Padhi (ZEUS)

- Its main purpose is to test basic functionalities of the Grid middleware
- It is implemented in collaboration with the Queen Mary University London (**Dave Kant**) that also takes part in the UK GridPP and LCG projects
- An example HEP application could be *Monte Carlo simulation*

---

**Grid Testbed: Technicalities**

- The Grid Testbed exploits 11 Linux PCs donated by H1 and HERA-B
- grid001 – grid011 run DESY Linux 4 (based on S.u.S.E. 7.2)
- Grid middleware of EDG 1.4 ([http://www.eu-datagrid.org](http://www.eu-datagrid.org))
- EDG is based on Globus 2.4 ([http://www.globus.org](http://www.globus.org))
- LCG-2 will use EDG 2
- EDG and Globus are built on/off RedHat Linux 6.2
- Normally Grid nodes are set up from an installation server (lcfg)
- The distribution allows to download binaries
- Those binaries where modified and installed on the DESY machines
Grid Testbed: Set-up

Authentication:
- Grid Security Infrastructure (GSI) based on PKI (openSSL)
- Globus Gatekeeper, Proxy renewal service

Grid Information Service (GIS):
- Grid Resource Information Service (GRIS)
- Grid Information Index Service (GIIS)

Resource Management:
- Resource Broker, Job Manager, Job Submission, Batch System (PBS), Logging and Bookkeeping

Storage Management:
- Replica Catalogue, GSI-enabled FTP, GDMP
- Replica Location Service (RLS)

Grid Testbed: Set-up cont’d

Testbed hardware:
- Mapping of services to logical and physical nodes.
- User Interface (UI)
- Computing Element (CE)
- Worker Node (WN)
- Resource Broker (RB)
- Storage Element (SE)
- Replica Catalogue (RC)
- Information Service (IS)
Grid Testbed: Schema

Grid Testbed: Monitor
Grid Testbed: Monitor

Andreas Gellrich, DESY IT Group
Grids@DESY, 27 October 2003
Grid Testbed: Monitor

- Submission of a simple job into the Grid Testbed
- Job delivers hostname and date of worker node
- Requires a certified user

Grid Testbed: Job Example

- Submission of a simple job into the Grid Testbed
- Job delivers hostname and date of worker node
- Requires a certified user

- Grid Environment
- Job script/binary which will be executed
- Job description by JDL
- Job submission
- Job status request
- Job output retrieval
Start.

```
grid003> export GLOBUS_LOCATION=/opt/globus
grid003> . $GLOBUS_LOCATION/etc/globus-user-env.sh
grid003> export EDG_LOCATION=/opt/edg
grid003> . $EDG_LOCATION/etc/edg-user-env.sh
grid003> export EDG_WL_LOCATION=$EDG_LOCATION
grid003> export RC_CONFIG_FILE=$EDG_LOCATION/etc/rc.conf
grid003> export GDMP_CONFIG_FILE=$EDG_LOCATION/etc/gdmp.conf
```

Grid Testbed: Proxy

```
grid003> grid-proxy-init
Your identity: /O=GermanGrid/OU=DESY/CN=Andreas Gellrich
Enter GRID pass phrase for this identity:
Creating proxy ............................
Done
Your proxy is valid until Thu Oct 23 01:52:00 2003

grid003> grid-proxy-info -all
subject : /O=GermanGrid/OU=DESY/CN=Andreas Gellrich/CN=proxy
issuer : /O=GermanGrid/OU=DESY/CN=Andreas Gellrich
type : full
strength : 512 bits
timeleft : 11:59:48
```
Grid Testbed: JDL

grid003> less script.sh
#!/usr/bin/zsh
host=`/bin/hostname`
date=`/bin/date`
echo "$host: $date"

grid003> less hostname.jdl
Executable  = "script.sh";
Arguments   = " ";
StdOutput   = "hostname.out";
StdError    = "hostname.err";
InputSandbox = ("script.sh");
OutputSandbox = ("hostname.out","hostname.err");
Rank        = other.MaxCpuTime;

Grid Testbed: Job Match

grid003> dg-job-listmatch hostname.jdl
Connecting to host grid006.desy.de, port 7771

***************************************************************************
COMPUTING ELEMENT IDs LIST
The following CE(s) matching your job requirements have been found:
- grid007.desy.de:2119/jobmanager-pbs-short
- grid007.desy.de:2119/jobmanager-pbs-long
- grid007.desy.de:2119/jobmanager-pbs-medium
- grid007.desy.de:2119/jobmanager-pbs-infinite
***************************************************************************
grid003> dg-job-submit hostname.jdl

Connecting to host grid006.desy.de, port 7771
Logging to host grid006.desy.de, port 15830

******************************************************************************

JOB SUBMIT OUTCOME
The job has been successfully submitted to the Resource Broker.
Use dg-job-status command to check job current status. Your job identifier (dg_jobId) is:
******************************************************************************

grid003> dg-job-status

Retrieving Information from LB server https://grid006.desy.de:7846
Please wait: this operation could take some seconds.

******************************************************************************

BOOKKEEPING INFORMATION:
Printing status info for the Job:

To be continued …
... continued:

---

Status = Scheduled
Last Update Time (UTC) = Thu Oct 23 13:47:38 2003
Job Destination = grid007.desy.de:2119/jobmanager-pbs-infinite
Status Reason = initial
Job Owner = /O=GermanGrid/OU=DESY/CN=Andreas Gellrich
Status Enter Time (UTC) = Thu Oct 23 13:47:38 2003
Location = GlobusJobmanager/grid007
***********************************************************************

---

Some bookkeeping information has not reached the LB server yet. Missing information should come from GlobusJobmanager
---

Status = Done
Job Destination = grid007.desy.de:2119/jobmanager-pbs-infinite
Status Reason = terminated
Job Owner = /O=GermanGrid/OU=DESY/CN=Andreas Gellrich
Status Enter Time (UTC) = Thu Oct 23 13:48:09 2003
***********************************************************************
... continued:

---

Status = OutputReady
Job Destination = grid007.desy.de:2119/jobmanager-pbs-infinite
Status Reason = terminated
Job Owner = /O=GermanGrid/OU=DESY/CN=Andreas Gellrich

Grid Testbed: Output

grid003> dg-job-get-output

*************************************************************************************************

JOB GET OUTPUT OUTCOME
Output sandbox files for the job:
have been successfully retrieved and stored in the directory:
/tmp/134721208077529

*************************************************************************************************
Grid Testbed: Result

```
grid003> ls -l /tmp/134721208077529
  total 4
  -rw-r--r-- 1 gellrich it 0 Oct 23 15:49 hostname.err
  -rw-r--r-- 1 gellrich it 39 Oct 23 15:49 hostname.out
```

```
grid003> less /tmp/134721208077529/hostname.out
```

Done!

---

Grid Testbed: Using RC

```
grid002> less file.jdl
InputData = {"LF:file.txt");
DataAccessProtocol = {"file", "gridftp");
ReplicaCatalog = ldap://grid001.desy.de:9011/
  lc=DesyCollection, rc=DesyRC, dc=grid001, dc=desy, dc=de";
```

```
grid002> ls -l /flatfiles/grid002/desy
-rwxr-xr-x 1 gdmp root 13 Oct 24 11:18 file.txt
```
grid003> edg-replica-manager-registerEntry -l file.txt -s
grid002.desy.de/flatfiles/grid002/desy/file.txt
configuration file: /opt/edg/etc/rc.conf
logical file name: file.txt
source: grid002.desy.de/flatfiles/grid002/desy/file.txt
protocol: gsiftp
SASL/GSI-GSSAPI authentication started
SASL SSF: 56
SASL installing layers
The program was successfully executed.

grid003> edg-replica-manager-ls
2000_nom.simrec.run file.txt ginit_copy sesion.txt test.py file ginit nic.txt setup

grid003> ldapsearch -x -H ldap://grid001.desy.de:9011 -b 'rc=DesyRC,dc=grid001,dc=desy,dc=de'

For LFN file.txt the following replicas have been found:
    location 1: grid002.desy.de/flatfiles/grid002/desy/file.txt
The program was successfully executed.
**Typical HEP job:**
- Executable, environment, input data, output data
- Monte Carlo simulation: big input, big output
- Analysis: huge input, small output
- Reprocessing: huge input, huge output

**RB problem:**
- Can not transfer huge data files with job
- Optimize for Computing resources or data resources

**LCG:**
- Let RB locate data and send the job there
- Let the WN retrieve data if necessary (not nice!)

---

**Grid Testbed: Minimal Set-up**

**Per site:**
- User Interface (UI) to submit jobs
- Computing Element (CE) to run jobs
- Worker Node (WN) to do the work
- Storage Element (SE) to provide data files
- [Grid Information Index Service (GIIS) as site-MDS]

**Per Grid:**
- Resource Broker (RB)
- Replica Catalogue (RC)
- Information Service (IS)
Grid Testbed: Observations

**System aspects:**
- DESY Linux 4 installation possible but clumsy
- Grid middleware penetrates system software

**Information Service:**
- Weak point in Globus (LDAP not scalable w/ writing)
- Rebuilt in EDG and LCG (R-GMA: RDBS rather than LDAP)

**Resource Broker:**
- Plays central role
- Needs to store all data coming with the job (resource intensive)

**Computing/Storage Element:**
- Probably scalable w/ more worker nodes (farms)
- Not fully understood yet

Grid Testbed: Future Plans

**Testbed:** (still under way)
- Get the missing pieces running (site-GIIS, SE)
- Connect again to QMUL
- Get more minimal sites involved: DESY Zeuthen, U Dortmund, ...
- The Grid Testbed is not application or experiment specific!
- The Grid Testbed is not meant for production!

**Prototyping of HEP applications:** (next step)
- Not studied in detail yet
- Open for various HEP applications
- Monte Carlo simulations is a good candidate

**Production:**
- Would need new (or) appropriate hardware
- Will there be a DESY-wide Grid infrastructure?
The dCache project is a joint effort between the Fermi National Laboratories and DESY, providing a fast and scalable disk cache system with optional Hierarchical Storage Manager connection.

Beside various other access methods, the dCache implements a Grid mass storage fabric.

It supports gsiftp and http as transport protocols and the Storage Resource Manager protocol (SRM) as Replica Management negotiation vehicle.

The system is in production for all DESY experiments, the CDF experiment at FNAL and for the CMS-US division at various locations, including CERN.

Grid Storage Fabric Abstraction

- **GRID Middleware**
- **SRM Client**
- **Data Transfer**
  - gsiftp
  - (ssl) HTTP
  - SRM

- **Storage Element**
- **Virtual Storage Layer**

- **dCache.org**
- **Desy**
- **Grid Karlsruhe**
- **Fermi**
- **San Diego**
- **CERN**
- **Jefferson**
- **QSM**
- **Tan**
- **Essere**
- **HPSS**
- **Castor**
- **Jasmine**
• http://www.lqcd.org/ildg/
• ILDG develops an international data grid for the lattice field theory community
• It develops an XML Schema suitable for describing the data generated by lattice field-theory
• A heterogeneous Grid-Of-Grids is needed, so a Web Services API will be used to connect existing archives (NERSC, CP-PACS, UKQCD)
• Experiments to test interoperability: aggregate replica catalogues, starting with just listing the contents of the Grid-of-Grids
• Allow public access via Web Services and/or HTTP
EGEE (Enabling Grids for E-Science in Europe) aims to integrate current national, regional and thematic computing and data grids to create a European Grid-empowered infrastructure for the support of the European Research Area, exploiting unique expertise generated by previous EU projects (DataGrid, Crossgrid, DataTAG, etc) and national Grid initiatives (UK e-Science, INFN Grid, Nordugrid, US Trillium, etc).

The EGEE consortium involves 70 leading institutions in 27 countries, federated in regional grids, with a combined capacity of over 20000 CPUs, the largest international Grid infrastructure ever assembled.

The EGEE vision is to provide distributed European research communities with a common market of computing, offering round-the-clock access to major computing resources, independent of geographic location, building on the EU Research Network Geant and NRNs. EGEE will support common Grid computing needs, integrate the computing infrastructures of these communities and agree on common access policies. The resulting infrastructure will surpass the capabilities of localized clusters and individual supercomputing centres, providing a unique tool for collaborative computer and data intensive e-Science. EGEE will work to provide interoperability with other major Grid initiatives such as the US NSF Cyberinfrastructure, establishing a worldwide Grid infrastructure.

EGEE is a two-year project in a four-year programme. Major implementation milestones after two years will provide the basis for assessing subsequent objectives and funding needs. Two pilot applications areas have been selected to guide the implementation and certify the performance of EGEE: the Particle Physics LHC Grid (LCG), where the computing model is based exclusively on a Grid infrastructure to store and analyze petabytes of data from experiments at CERN; Biomedical Grids, where several communities are facing equally daunting challenges to cope with the flood of bioinformatics and healthcare data, such as the proposed HealthGrid association.

The project objectives will be achieved by the aggregation of the human and computing resources of regional Grid federations established by EGEE, by a complete re-engineering of the middleware, and by a pro-active program of outreach and training to attract and support the widest possible variety of scientific communities in the ERA.

Three-fold mission:

- To deliver production level Grid services
- To carry out a professional Grid middleware re-engineering
- To ensure an outreach and training effort
Layer 0: Resource Centers (RC)
DESY

Layer I: Regional Operations Centres (ROC)
FZ Karlsruhe

Layer II: Core Infrastructure Centres (CIC)
FZ Karlsruhe

Layer III: Operations Management Centre (OMC)
CERN

Networking (Network Activities, NA):
• NA1: Management of the I3
• NA2: Dissemination and Outreach
• NA3: Training and Induction
• NA4: Application identification and Support
• NA5: Policy and International Cooperation

Services (Specific Service Activities, SA):
• SA1: Grid operations, Support and Management
• SA2: Network resource provision

Middleware (Joint Research Activities, JRA):
• JRA1: Middleware Engineering and Integration Research
• JRA2: Quality Assurance
• JRA3: Security
• JRA4: Network Service Development
EGEE@DESY

- Volker Gülzow (Member)
- Andreas Gellrich (Deputy)
- SA1 (Specific Service Activities)
- European Grid Support, Operation and Management
- Partner of the German ROC
- Support Centre and Resource Centre
- User and Administrator Support, CPU and Storage provision

- DESY will receive funding of ~148 k€ for 2 years
- DESY is committed to provide infrastructure

EGEE: Timeline

- May 2003: Proposal
- 10 Oct 2003: Submission of Technical Annex (TA)
- 17 Oct 2003: Final negotiation with EU
- 20 Oct 2003: DESY signs contract (CPF)
- 30 Nov 2003: Consortium Agreement (CA) signing
- 1 Apr 2004: Project start
- 16-23 April 2004: Kick-off Meeting, Ireland
What I missed

- GAN (Global Accelerator Networks)
- Other Toolkit Approaches
- Other LCG middleware projects (VDT,...)
- Non-HEP Grid applications
- Commercial / industrial applications
- Many political aspects
- Many, many technical details
- ...

Conclusions

- The Grid has developed from a smart idea to a real implementation
- Grid infrastructure will be standard in (e)-science in the future
- LHC can not live w/o Grids
- DESY could live w/o Grids now, but would miss the future
- There are projects under way at DESY
- See: http://www-it.desy.de/physics/projects/grid/
- The EGEE project opens the Grid field for DESY