High Performance & Grid Computing in Europe

Dieter Kranzlmüller
Ludwig-Maximilians-Universität (LMU) & Leibniz Supercomputing Centre (LRZ)
Munich, Germany
The Leibniz Supercomputing Centre
of the Bavarian Academy of Sciences and Humanities

• Computing Centre for all Munich Universities
• Regional Computing Centre for all Bavarian Universities
• National Supercomputing Centre of Germany

By Ernst A. Graf

© 2009 Leibniz Supercomputing Centre
HLRB-II: SGI Altix 4700

9728 Cores
62.3 TFlop/s Peak
39 TByte Memory
103 metric tons
~1100 kVA

Foto Helmut Payer, produced by gsiCom
HPC & Grid Resources

4th IGSSE Forum,
Raitenhaslach
HPC & Grid Ecosystem

- Computational Fluid Dynamics: Optimisation of turbines and wings, noise reduction, air conditioning in trains
- Fusion: Plasma in a future fusion reactor (ITER)
- Astrophysics: Origin and evolution of stars and galaxies
- Solid State Physics: Superconductivity, surface properties
- Geophysics: Earthquake scenarios
- Material Science: Semiconductors
- Chemistry: Catalytic reactions
- Medicine and Medical Engineering: Blood flow, aneurysms, air conditioning of operating theatres
- Biophysics: Properties of viruses, genome analysis
- Climate research: Currents in oceans
Examples of Applications
The ESFRI Vision for a European HPC service

- European HPC-facilities at the top of an HPC provisioning pyramid
  - Tier-0: 3-5 European Centres
  - Tier-1: National Centres
  - Tier-2: Regional/University Centres

- Creation of a European HPC ecosystem involving all stakeholders
  - HPC service providers on all tiers
  - Grid Infrastructures
  - Scientific and industrial user communities
  - The European HPC hard- and software industry
Europe prepares the creation of a persistent pan-European HPC service, consisting of several tier-0 centres providing European researchers with access to capability computers and forming the top level of the European HPC ecosystem. PRACE is a project funded in part by the EU’s 7th Framework Programme.

### PRACE Partners

**Germany**  
EPSRC - Engineering and Physical Sciences Research Council,

**UK**  
GENCI - Grand Equipement national pour le Calcul Intensif

**France**  
BSC - Barcelona Supercomputing Center

**Spain**  
NCF - Netherlands Computing Facilities Foundation

**Netherlands**  
EPSRC - Engineering and Physical Sciences Research Council,

**Finland**  
GRNET - Greek Research and Technology Network

**Greece**  
CINECA - Consorzio Interuniversitario per ilCalcolo Automatic dell'Italia Nord Orientale

**Italy**  
UNINETT Sigma AS, sigma.uninett.no

**Norway**  
GUP - Institut für Informatik der Johannes Kepler Universität Linz

**Austria**  
Poznan Supercomputing and Networking Center

**Poland**  
Universidade de Coimbra

**Portugal**  
KTH - Kungl Tekniska Högskolan

**Sweden**  
CSCS - Swiss National Supercomputing Centre

**Switzerland**  
Informatics Institute, ITU, www.be.itu.edu.tr

**Turkey**  
Barcelona Supercomputing Center

**United Kingdom**  
Nordic HPC Centre

**Germany**  
Nordic HPC Centre

**Spain**  
Nordic HPC Centre

**Netherlands**  
Nordic HPC Centre

**Finland**  
Nordic HPC Centre

**Greece**  
Nordic HPC Centre

**Italy**  
Nordic HPC Centre

**Norway**  
Nordic HPC Centre

**Austria**  
Nordic HPC Centre

**Poland**  
Nordic HPC Centre

**Portugal**  
Nordic HPC Centre

**Sweden**  
Nordic HPC Centre

**Switzerland**  
Nordic HPC Centre

**Turkey**  
Nordic HPC Centre
PRACE Objectives in a Nutshell

• Provide world-class systems for world-class science
• Create a single European (legal) entity
• Deploy 3 – 5 systems of the highest performance level (tier-0)
• Ensure diversity of architectures
• Provide support and training
LRZ Extension Buildings

Fotomontage

Südansicht

Nordansicht
Future Petaflop/s Computer Technologies beyond 2010

• Evaluation of emerging multi-petascale-technology following the requirements of HPC users
• Implementation of a strategy that guarantees a continuous HPC technology evaluation and system evolution within the PRACE Research Infrastructure
• Fostering the development of components for future multi-petascale production systems in cooperation with European and international HPC industry
Performance and Memory

10-Fold every 4 years
Double every 14.5 Month

HLRB2: SGI Altix 4700
Linux-Cluster

HLRB1: Hitachi SR8000
Linux-Cluster

IBM SP2
Cray T90

Fujitsu VPP

KSR

Cray Y-MP

Gigaflops, GByte

10,000,000
1,000,000
100,000
10,000
1,000
100
10
1


GPlofs GByte
Number of Cores

HP-CAST 12, Madrid Spain

LRZ, High Performance Computing Group, Matthias Brehm, June 10
Energy Consumption

Energy Cost 2009: 0.13 € / kWh
Today: PUE ~1.4
Ergebnisse (LLNL)

SMG2000, Semicoarsening
Multigrid Solver
ASCI Purple Benchmark Suite
MD Simulation

- NAMD/VMD
- Interactive MD Simulation
- 1 Mio Atoms
- 2000 Cores on HLRB-II
- Displayed in 3D stereo in Dresden
- Ferdinand Jamitzky & Helmut Satzger
LRZ Extension Buildings

Fotomontage

Südansicht

Nordansicht
HPC & Grid Ecosystem

HP-CAST 12, Madrid Spain
4th IGSSE Forum, Raitenhaslach


Archeology
Astronomy
Astrophysics
Civil Protection
Comp. Chemistry
Earth Sciences
Finance
Fusion
Geophysics
High Energy Physics
Life Sciences
Multimedia
Material Sciences
...

> 260 Sites
> 55 Countries
> 150,000 CPUs
> 28 PetaBytes Disk
> 14,000 Users
> 200 VOs
> 330,000 Jobs/Day

Large Hadron Collider (LHC)
LHC Detektoren

40 MHz (40 TB/sec)

level 1 - special hardware

75 KHz (75 GB/sec)

level 2 - embedded processors

5 KHz (5 GB/sec)

level 3 - PCs

100 Hz (100 MB/sec)

data recording & offline analysis

4th IGSSE Forum, Raitenhaslach
Biomedicine: WISDOM

- World-wide In Silico Docking On Malaria
  - Neglected diseases
  - Biomedical data challenges
  - Docking experiments:
    - 46 million ligands
  - Runtime: 6 weeks
  - Data: 1 TByte
  - 1000 CPUs in 15 countries
    (~80 CPU years)
Example: Biomedical Simulation

Delete

angle measurement:
base: (1.08 1.86 -0.28) meters
direction 1: (1.22 1.81 -0.40) meters
direction 2: (1.22 1.93 -0.39) meters
angle: 38.32 degree
Blood Flow Simulation

- Parallel simulation on the grid
- Online visualization at desktop
- „Invisible Grid“
g-Eclipse
Flooding Crisis Management
Evolution

National

Global

Testbeds

Routine Usage

Utility Service
Grids in Europe

www.eu-egi.eu
European Grid Initiative

Objectives:
• Ensure the long-term sustainability of the European e-infrastructure
• Coordinate the integration and interaction between National Grid Infrastructures
• Operate the European level of the production Grid infrastructure for a wide range of scientific disciplines to link National Grid Infrastructures

EGI Grid Infrastructure should be
• a large-scale, production Grid infrastructure
• built on national grids that interoperate seamlessly at many levels,
• offering reliable and predictable services to a wide range of applications
Status EGI

EGI-InSPIRE: Integrated Sustainable Pan-European Infrastructure for Researchers in Europe

- Duration: 4 years (starting 1 May 2010)
- Costs: 70 MEUR, incl. 25 MEUR EU
- Total costs incl. NGI resources: ca. 330 MEUR
- Partner: EGI.eu, 37 NGIs, 2 EIROs

EGI.eu: EGI organization

- Location: Amsterdam, The Netherlands
- Status approved, EGI.eu Executive Board elected
EGI Operations Tasks

Operation of tools and services
- Grid configuration repositories
- Grid accounting repositories
- Grid repositories for SLA compliance and performance monitoring
- Grid operations portal
- NGI Grid oversight

Security
- Security policy development and maintenance
- Coordination of security and incident response
- Expert team for security vulnerabilities

User support
- Central ticket handling system
- Gathering requirements for user support tools

Other international tasks
- MW deployment/roll-out and support
- Resource allocation & brokering support
- Interoperations between NGI’s and with other grids
- Network support

EGI Blueprint Proposal (V3.0)
- Functions of EGI
- Financing of EGI
- Transition to EGI

4th IGSSE Forum, Raitenhaslach
EGI Infrastructure

International Scientific and Research Collaboration

EGI.eu

European-level Grid Services

- gLite
- UNICORE
- ARC
- Middleware A
- Middleware B
- Middleware B
- Middleware X

National Grid Initiative 1
National Grid Initiative 2
...
National Grid Initiative N
EGI Operations = EGI.eu critical services + NGI international tasks
**ComputingService (DB)**

<table>
<thead>
<tr>
<th>service</th>
<th>type</th>
<th>domainID</th>
<th>informationProvider</th>
<th>sourceAddr</th>
<th>insertTime</th>
</tr>
</thead>
<tbody>
<tr>
<td>UNICORE 6 Target System Factory Service</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>alnitak.ari.uni-heidelberg.de</td>
<td>org.teragrid.ws-gram AstroGrid-D</td>
<td>MDS4</td>
<td>129.187.254.39</td>
<td>2008-08-07 11:25:16</td>
<td></td>
</tr>
<tr>
<td>anticyclone.dkrz.de</td>
<td>org.teragrid.ws-gram dkrz.de</td>
<td>MDS4</td>
<td>129.187.254.39</td>
<td>2008-08-07 11:25:16</td>
<td></td>
</tr>
<tr>
<td>aprilia.izb.uni-leipzig.de</td>
<td>org.teragrid.ws-gram</td>
<td>MDS4</td>
<td>129.187.254.39</td>
<td>2008-08-07 11:25:16</td>
<td></td>
</tr>
<tr>
<td>arminius-grid.uni-paderborn.de</td>
<td>org.teragrid.ws-gram PC2</td>
<td>MDS4</td>
<td>129.187.254.39</td>
<td>2008-08-07 11:25:16</td>
<td></td>
</tr>
<tr>
<td>astar.aip.de</td>
<td>org.teragrid.ws-gram AstroGrid-D</td>
<td>MDS4</td>
<td>129.187.254.39</td>
<td>2008-08-07 11:25:16</td>
<td></td>
</tr>
<tr>
<td>buran.aei.mpIng.de</td>
<td>org.teragrid.ws-gram AstroGrid-D</td>
<td>MDS4</td>
<td>129.187.254.39</td>
<td>2008-08-07 11:25:16</td>
<td></td>
</tr>
<tr>
<td>c3grid.rz.uni-koeln.de</td>
<td>org.teragrid.ws-gram de.uni-koeln</td>
<td>MDS4</td>
<td>129.187.254.39</td>
<td>2008-08-07 11:25:16</td>
<td></td>
</tr>
<tr>
<td>cashmere.aip.de</td>
<td>org.teragrid.ws-gram AstroGrid-D</td>
<td>MDS4</td>
<td>129.187.254.39</td>
<td>2008-08-07 11:25:16</td>
<td></td>
</tr>
<tr>
<td>ce-1-fzk.gridka.de</td>
<td>pbspro-gLite3</td>
<td>FZK-LCG2</td>
<td>BDII</td>
<td>wrdmon-bdiif.zdk.de</td>
<td>2008-09-19 17:13:20</td>
</tr>
<tr>
<td>ce-2-fzk.gridka.de</td>
<td>pbspro-gLite3</td>
<td>FZK-LCG2</td>
<td>BDII</td>
<td>wrdmon-bdiif.zdk.de</td>
<td>2008-09-19 17:13:20</td>
</tr>
<tr>
<td>ce-3-fzk.gridka.de</td>
<td>pbspro-gLite3</td>
<td>FZK-LCG2</td>
<td>BDII</td>
<td>wrdmon-bdiif.zdk.de</td>
<td>2008-09-19 17:13:20</td>
</tr>
</tbody>
</table>

**VO-specific View (OGSA-DAI)**

SQLQuery: SELECT name, type, totalJobs, runningJobs, waitingJobs, informationProvider FROM ComputingService_VO VO: vi uk.org.oagsa-dai.resource.request.status.COMPLETED

<table>
<thead>
<tr>
<th>name</th>
<th>type</th>
<th>totalJobs</th>
<th>runningJobs</th>
<th>waitingJobs</th>
<th>informationProvider</th>
</tr>
</thead>
<tbody>
<tr>
<td>arminius-grid.uni-paderborn.de</td>
<td>org.teragrid.ws-gram</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>MDS4</td>
</tr>
<tr>
<td>dgiref-globus.fzk.de</td>
<td>org.teragrid.ws-gram</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>MDS4</td>
</tr>
<tr>
<td>dgrid-glite1.rz.rwth-aachen.de</td>
<td>org.teragrid.ws-gram</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>MDS4</td>
</tr>
<tr>
<td>dmon-unic.rz-juelich.de$9115$FZJ-JUGGLE-DMON</td>
<td>org.teragrid.ws-gram</td>
<td>151</td>
<td>151</td>
<td>0</td>
<td>MDS4</td>
</tr>
<tr>
<td>udo-gt01.grid.uni-dortmund.de</td>
<td>org.teragrid.ws-gram</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>MDS4</td>
</tr>
<tr>
<td>udo-gt03.grid.tu-dortmund.de</td>
<td>org.teragrid.ws-gram</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>MDS4</td>
</tr>
</tbody>
</table>
Generic VO-Model

Formation
Identification > Initialization > Operation > Adaptation > Termination

Community-Layer

VOA

VOB

VO-Layer

RO-Layer

RO1

RO2

S

R

S

R

S

R

S

R

39

4th IGSSE Forum,
Raitenhaslach
VO Management?

- Lifecycle management
- Membership management
- Resource management
- Security management
- Constraints
  - Heterogeneity
  - No central control
  - Autonomous administrative domains
  - Management on attribute level
Cloud Computing

“Cloud Computing is on-demand access to virtualized IT resources that are sourced inside or outside of a data center, scalable, shared by others, simple to use, paid for via subscription or as you go, and accessible over the Web.”

[Dr. Behrend Freese (Zimory GmbH)]

“Cloud Services = Consumer and Business products, services and solutions that are delivered and consumed in realtime over the Internet”

[IDC – Analyze the Future]

“A Cloud is a pool of abstracted, highly scalable and managed IT infrastructure that provides customers’ applications and will be charged by actual use.“

[Forrester Research]
4th IGSSE Forum,
Raitenhaslach
Grids vs. Clouds

**Grids**
- Sciences, research, collaboration
- Collaborative use of heterogeneous resources
- Local autonomy
- No centralized management
- Publicly funded and operated
- No business model
- Complex interfaces
- Open standards (OGF, …)
- Interoperability

**Clouds**
- Business and web-based applications
- Homogeneous and virtualized resources
- One owner
- Centralized management
- Business intentions
- Pay-per-use business model
- Simple interfaces
- No standards as of today
- No interoperability (vendor lock-in)
Experimental results: main memory

• Virtualization of main memory with high performance
• Caches may contaminate!

Testing conditions

- Benchmark: RAMspeed
- Test: Reading
- Modus: sequentiel
- # VMs: 1
- RAM: 1 GB
- vCPUs: 1
Benchmarking Amazon Cloud

Benchmarking Amazon Cloud


**Figure 3.** MPI Bandwidth Performance in the MPTEST Benchmark on the NCSA and EC2 Clusters

**Figure 4.** MPI Latency Performance in the MPTEST Benchmark on the NCSA and EC2 Clusters

4th IGSSE Forum, Raitenhaslach
Pricing: Cloud vs. Hosting

1 CPU-Unit, 1,7 GB Memory, 100 GB Disk
Network Traffic: IN 50 GB, OUT 100 GB

Quelle: Zimory GmbH

4th IGSSE Forum, Raitenhaslach
Pricing: Cloud vs. Hosting

1 CPU-Unit, 1,7 GB Memory, 200 GB Disk
Network Traffic: IN 100 GB, OUT 300 GB

Quelle: Zimory GmbH

4th IGSSE Forum, Raitenhaslach
Summary

• European HPC & Grid Activities:
  – PRACE
    = Partnership for Advanced Computing in Europe
  – EGI
    = European Grid Initiative
• Integration into a common, persistent e-infrastructure for science and research
• To support international scientific collaboration
High Performance & Grid Computing in Europe

Dieter Kranzlmüller
kranzlmueller@ifi.lmu.de