

EGEE Enabling Grids for E-science

Grid Applications and Use Cases

F Dr. Rüdiger Berlich,
Forschungszentrum Karlsruhe / Germany
Mumbai, 11.02.06

Slides contributed by FZK + EGEE Team

www.eu-egee.org

School@chep06

Information Society

INFSO-RI-508833

EGEE Enabling Grids for E-science

Exponential Growth ...

Triumph of Light – *Scientific American*. George Stix, January 2001

School@chep06

INFSO-RI-508833

Grid Applications and Use Cases, February 11, 2006

EGEE Enabling Grids for E-science

Evolution

„When the network is as fast as the computer's internal links, the machine dis-integrates across the net into a set of special purpose appliances“
(Gilder Technology Report, June 2001)

School@chep06

INFSO-RI-508833

Grid Applications and Use Cases, February 11, 2006

EGEE Enabling Grids for E-science

Technical Limitations

- „Speed“ of a network consists of two components
- Bandwidth (scales to any number)
- Latency (doesn't scale)
- Possible application types in a distributed environment are limited by latency

School@chep06

INFSO-RI-508833

Grid Applications and Use Cases, February 11, 2006

EGEE Enabling Grids for E-science

The GRID vision

Researchers perform their activities regardless geographical location, interact with colleagues, share and access data

The Grid: networked data processing centres and "middleware" software as the "glue" of resources.

Scientific instruments and experiments provide huge amount of data

School@chep06

INFISO-RI-508833 Grid Applications and Use Cases, February 11, 2006 5

EGEE Enabling Grids for E-science

Use Case in Science: MAGIC

- **Ground based Air Cerenkov Telescope 17 m diameter**
- **Physics Goals:**
 - Origin of VHE Gamma rays
 - Active Galactic Nuclei
 - Supernova Remnants
 - Unidentified EGRET sources
 - Gamma Ray Burst
- **MAGIC II will come 2007**
- **Grid added value**
 - Enable "(e-)scientific" collaboration between partners
 - Enable the cooperation between different experiments
 - Enable the participation on Virtual Observatories

School@chep06

INFISO-RI-508833 Grid Applications and Use Cases, February 11, 2006 6

EGEE Enabling Grids for E-science

Use Case in Science: Bioinformatics

- **GPS@: Grid Protein Sequence Analysis**
 - **Gridified version of NPSA web portal**
 - Offering proteins databases and sequence analysis algorithms to the bioinformaticians (3000 hits per day)
 - Need for large databases and big number of short jobs
 - **Objective:** increased computing power
 - **Status:** 9 bioinformatic softwares gridified
 - **Grid added value:** open to a wider community with larger bioinformatic computations
- **xmipp_MLrefine**
 - **3D structure analysis of macromolecules**
 - From (very noisy) electron microscopy images
 - Maximum likelihood approach to find the optimal model
 - **Objective:** study molecule interaction and chem. properties
 - **Status:** algorithm being optimised and ported to 3D
 - **Grid added value:** parallel computation on different resources of independent jobs

School@chep06

INFISO-RI-508833 Grid Applications and Use Cases, February 11, 2006 7

EGEE Enabling Grids for E-science

Bio-medicine applications

September 21st, 2001 *Datagrid Meeting, Lyon* 15

Mammogrid -> AliEn

School@chep06

INFISO-RI-508833 Grid Applications and Use Cases, February 11, 2006 8

EGEE Enabling Grids for E-science

Similarity search

Similarity computation

Job	Status	Target
2142	Completed	bioinformatics
2143	Completed	bioinformatics
2144	Completed	bioinformatics
2145	Completed	bioinformatics
2146	Completed	bioinformatics
2147	Completed	bioinformatics
2148	Completed	bioinformatics
2149	Completed	bioinformatics
2150	Completed	bioinformatics
2151	Completed	bioinformatics
2152	Completed	bioinformatics
2153	Completed	bioinformatics
2154	Completed	bioinformatics
2155	Completed	bioinformatics
2156	Completed	bioinformatics
2157	Completed	bioinformatics
2158	Completed	bioinformatics
2159	Completed	bioinformatics
2160	Completed	bioinformatics
2161	Completed	bioinformatics
2162	Completed	bioinformatics
2163	Completed	bioinformatics
2164	Completed	bioinformatics
2165	Completed	bioinformatics
2166	Completed	bioinformatics
2167	Completed	bioinformatics
2168	Completed	bioinformatics
2169	Completed	bioinformatics
2170	Completed	bioinformatics
2171	Completed	bioinformatics
2172	Completed	bioinformatics
2173	Completed	bioinformatics
2174	Completed	bioinformatics
2175	Completed	bioinformatics
2176	Completed	bioinformatics
2177	Completed	bioinformatics
2178	Completed	bioinformatics
2179	Completed	bioinformatics
2180	Completed	bioinformatics
2181	Completed	bioinformatics
2182	Completed	bioinformatics
2183	Completed	bioinformatics
2184	Completed	bioinformatics
2185	Completed	bioinformatics
2186	Completed	bioinformatics
2187	Completed	bioinformatics
2188	Completed	bioinformatics
2189	Completed	bioinformatics
2190	Completed	bioinformatics
2191	Completed	bioinformatics
2192	Completed	bioinformatics
2193	Completed	bioinformatics
2194	Completed	bioinformatics
2195	Completed	bioinformatics
2196	Completed	bioinformatics
2197	Completed	bioinformatics
2198	Completed	bioinformatics
2199	Completed	bioinformatics
2200	Completed	bioinformatics

Job monitoring Results visualization images

Ranked list of

Source image Most similar images Low score images

School *chep06*

INFSO-RI-508833 Grid Applications and Use Cases, February 11, 2006 9

EGEE Enabling Grids for E-science

Bio-medicine applications

3.3 Heart Modeling

- Objectives: modeling heart anatomy, dynamics and physiology for heart image processing
 - bio-mechanical model
 - electrical model
 - very complex structure
 - biological scale out of range
- Finite Element modeling
 - elements oriented in heart fibers direction: fine resolution
 - electrical propagation model based on bidomain theory
 - 4D model (3D+T)

Bio engineering research group, Auckland

School *chep06*

GGF1 - DataGrid WP 10 - March 2006

INFSO-RI-508833 Grid Applications and Use Cases, February 11, 2006 1

EGEE Enabling Grids for E-science

Earth observation applications

ENVIAT

- 3500 Meuro programme cost
- Launched on February 28, 2002
- 10 instruments on board
- 200 Mbps data rate to ground
- 400 Tbytes data archived/year
- ~100 'standard' products
- 10+ dedicated facilities in Europe
- ~700 approved science user projects

School *chep06*

INFSO-RI-508833 Grid Applications and Use Cases, February 11, 2006 1

EGEE Enabling Grids for E-science

Virtual Observatory

School *chep06*

INFSO-RI-508833 Grid Applications and Use Cases, February 11, 2006 1

EGEE Enabling Grids for E-science

Flood simulation

School chep06

Sample Vah river

Computer vision

Geographical Information Systems

Results: flow + water depths

INFSO-RI-508833 Grid Applications and Use Cases, February 11, 2006 1

EGEE Enabling Grids for E-science

Engineering applications

Network for Earthquake Engineering Simulation

- NEESgrid: national infrastructure to couple earthquake engineers with experimental facilities, databases, computers, & each other
- On-demand access to experiments, data streams, computing, archives, collaboration

NEESgrid: Argonne, Michigan, NCSA, UIUC, USC

School chep06

INFSO-RI-508833 Grid Applications and Use Cases, February 11, 2006 1

EGEE Enabling Grids for E-science

Grid Applications: Art

Paintings are being scanned in at 30 GB each in the EU CRISATEL Project

Museo Virtual de Artes El Pais (MUVA)

Books are being scanned in at 767 MB per page 1/2 Terabyte for Gutenberg Bible

School chep06

INFSO-RI-508833 Grid Applications and Use Cases, February 11, 2006 1

EGEE Enabling Grids for E-science

Connecting People: Access Grid

Remote video

Visualisation

Microphones

Cameras

School chep06

INFSO-RI-508833 Grid Applications and Use Cases, February 11, 2006 1

EGEE Enabling Grids for E-science

Summary / Conclusion

- Many different Grid types (although not in the mainstream)
- Typical Areas:
 - „Embarrassing Parallel“ Applications (or „nicely parallel“)
 - Collaboration
 - Distributed Databases
 - Some real parallel processing (rare)

School
chep06

INFISO-RI-508833 Grid Applications and Use Cases, February 11, 2006 1

EGEE Enabling Grids for E-science

Use Case: Business (1) – CPU Cycles

Welcome to the Sun Grid.
Simple, secure compute utilities for \$1.

- Selling compute cycles
- Sun: 1\$ / CPU-hour

IBM

School
chep06

INFISO-RI-508833 Grid Applications and Use Cases, February 11, 2006 1

EGEE Enabling Grids for E-science

Use Case: Business (2) – DAME

Slide taken from talk given at GridKa School by Dr. Tom Jackson, University of York

School
chep06

INFISO-RI-508833 Grid Applications and Use Cases, February 11, 2006 1

EGEE Enabling Grids for E-science

Use Case: Business (2) – DAME

DAME Objectives

- Building a demonstration system as proof of concept for Grid technology in the aerospace diagnostic domain.
- Three primary Grid challenges:
 - Management of large, distributed and heterogeneous data repositories;
 - Rapid data mining and analysis of fault data;
 - Information management and data fusion for diagnosis/prognosis applications;
- Other key (commercial) issues:
 - Remote, secure access to flight data and other operational data and resources;
 - Management of distributed users and resources;
 - Quality of Service issues (and Service Level Agreements)

Slide taken from talk given at GridKa School by Dr. Tom Jackson, University of York

School
chep06

INFISO-RI-508833 Grid Applications and Use Cases, February 11, 2006 2

I'd like to thank
the organisers of this school and the audience,
the German Federal Ministry of Education
and Research (BMB+F),
the EGEE project and its representatives
as well as Forschungszentrum Karlsruhe!

Questions, please !!



bmb+f - Förderschwerpunkt
Hadronen -
und Kernphysik
Großgeräte der physikalischen
Grundlagenforschung