HealthGrid and SHARE: retrospect and prospect for grids in health

Yannick Legré (HealthGrid) on behalf of the SHARE Consortium

http://www.eu-share.org
http://www.healthgrid.org

SHARE: Structuring and supporting Healthgrids Activities and Research in Europe

Slides credit: Tony Solomonides

The Share project (FP6-2005-IST-027694) is funded by the European Commission Information Society and Media DG under the 6th Framework Programme.

ISGC 2008 - Taipei - 7th - 11th April, 2008
EC Framework Programme 6 ‘Specific Support Action’ project
27 months, 1st January 2006 to 31st March 2008

with

- CNRS/IN2P3
- HealthGrid
- Universidad Politécnica de Valencia
- University of the West of England, Bristol
- Research Centre for Computer and Law (CRID) - University of Namur
- European Health Management Association
- Empirica GmbH

- Argonne National Laboratory
- Academia Sinica Grid Computing Centre
  - APAMI (Asia-Pacific Association for Medical Informatics)
SHARE Objectives

- **SHARE** to define milestones for
  - wide deployment and adoption of **healthgrids in Europe**
  - action plan for a **European e-Health Area**
- The project had to
  - assess the **status quo** and set **targets**
  - identify key **gaps, barriers and opportunities**
  - establish
    - short and long term objectives
    - key developments
    - actors to achieve the vision
Background

- The concept of “grids for health” was described in the HealthGrid White Paper in 2005. It set out a vision of the opportunities and potential benefits offered by applying grids in different areas of biomedicine and healthcare.

- The HealthGrid vision relies on the setting up of grid infrastructures for medical research, healthcare, and the life sciences.

- HealthGrid itself arose from a number of projects in grid applications to medicine and healthcare from about 2001 onwards. They spanned:
  - health informatics: screening, epidemiology, public health, etc.
  - clinical informatics: diagnostics, decision support, care planning, etc.

The SHARE project (2004-2007) is funded by the European Commission Information Society and Media DG under the 6th Framework Programme.
"Grid infrastructures for biomedical informatics" implies:
- the availability of grid services, most notably for data and knowledge management;
- the deployment of these services on infrastructures involving healthcare centres (e.g. hospitals), medical research laboratories and public health administrations; and
- the definition and adoption of international standards and interoperability mechanisms for medical information stored on the grid.

Biomedical informatics
- a concurrent development
- convergence and synergy between medical informatics and bioinformatics
- leading to two new approaches to medicine...
SHARE Objectives

- SHARE to define
  - what has to be done,
  - when - and in what sequence,
  - by whom,
  - and how?
- Turns out action required in several domains:
  - technical research and development
  - standards and security for real world deployment
  - squaring up to ethical and legal issues
  - community acceptance and economic investment
What is the goal?

An environment, created through the sharing of resources, in which heterogeneous and dispersed health data at different levels:

- molecular data (e.g. genomics, proteomics)
- cellular data (e.g. pathways)
- tissue data (e.g. cancer types, wound healing)
- personal data (e.g. EHR)
- population (e.g. epidemiology)

as well as applications, can be accessed by all users as a tailored information system according to their level of authorisation and without loss of quality of information or service.
Technical Challenges

- Distributed data integration and computing
  - Security
  - Performance
  - Usability

- Standards
  - Need for reference implementations of standard grid services
  - Bridge the gap between medical informatics standards and grid standards (e.g. grid-enabled DICOM)
  - Lack of standard open source ontologies in medical informatics

- Grid deployment in medical research centres
  - Easy installation of secure grid nodes

The Share project (FP6-2005-IST-027697) is funded by the European Commission Information Society and Media DG under the 6th Framework Programme.
Other challenges

- Specific features of the community
  - Patient ownership of her or his data
  - Hospitals IT policies vs grids
- Technology transfer between projects
- Development of best practices
  - Interfacing IT resources for clinical routine to grid
  - Data sharing (and major ethical implications)
- Raising awareness of grids
  - Need to build on success stories
HealthGrid ‘SOA’

The classic grid architecture assumed by SHARE

- Core services are generic; no medical or healthcare specialization assumed
- Healthgrid services are generic services (e.g. pseudonymization, image storage) and may be used by different special applications
- Domain-specific applications may require additional services (e.g. mammogram standardization); these may also be made generic.
Toward a roadmap

**Phase 1**
- Sustainable computing grid
  - Reference implementation of grid services

**Phase 2**
- Sustainable data grid
  - Reference distribution of grid services
- Sustainable knowledge grid
  - Agreed medical informatics & grid standards
- Generalized use of knowledge grids
  - Agreed open source medical ontologies

The SHARE project (FP6-2005-IST-027694) is funded by the European Commission Information Society and Media DG under the 6th Framework Programme.

ISGC 2008 - Taipei - 7th - 11th April, 2008
In the first phase:

- **GD.1** A sustainable computing grid infrastructure for the medical research community
- **IT.1** A reference implementation of grid services using standard web service technology and allowing computation and secure manipulation of distributed data
- **GD.2** A sustainable data grid for a well defined medical research topic
  - Distributed storage and distant query of medical data
- **IT.2** A reference distribution of a reference implementation of grid services for the installation of grid nodes in medical research centres
Milestones II

In the second phase:

- **IT.3** An agreed set of standards for sharing medical images and records on the grid
- **GD.3** A knowledge grid for a well defined medical research topic
  - Distributed data integration and computing
- **IT.4** Agreed and implemented open source medical ontologies
- **GD.4** Generalized use of knowledge grids
Computational Grids

Research challenges for:
- Computing grids
- Data grids
- Knowledge grids

On demand access

Interoperability of Infrastructures

User friendliness

Quality of service

TIME

RCCG1
RCCG2
RCCG3
RCCG4
RCCG5
RCCG6
RCCG7
RCCG8
RCCG9
RCCG10

The Share project (FP6-2005-1ST-027694) is funded by the European Commission Information Society and Media DG under the 6th Framework Programme.

ISGC 2008 - Taipei - 7th - 11th April, 2008
<table>
<thead>
<tr>
<th>Challenge</th>
<th>Community</th>
<th>Description of the requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCCG1</td>
<td>VPH</td>
<td>Access to grid resources on demand.</td>
</tr>
<tr>
<td>RCCG2</td>
<td>VPH</td>
<td>Transparent job submission to cluster and supercomputer grids. Easy transfer of tasks between grid infrastructures</td>
</tr>
<tr>
<td>RCCG3</td>
<td>VPH</td>
<td>Automatic migration of simulations between different scales.</td>
</tr>
<tr>
<td>RCCG4</td>
<td>VPH</td>
<td>User friendly access. Lower barrier to adoption.</td>
</tr>
<tr>
<td>RCCG5</td>
<td>VPH</td>
<td>Transparent access to different grids.</td>
</tr>
<tr>
<td>RCCG6</td>
<td>EPI</td>
<td>Need for real fault-tolerant scheduling systems.</td>
</tr>
<tr>
<td>RCCG7</td>
<td>EPI</td>
<td>Easily installed grid middleware for health environments. Low maintenance and administration.</td>
</tr>
<tr>
<td>RCCG8</td>
<td>EPI</td>
<td>Exploitation models and guaranteed QoS for services. Advance resource reservation with pre-negotiated QoS.</td>
</tr>
<tr>
<td>RCCG9</td>
<td>EPI</td>
<td>Need for scalable job scheduling system.</td>
</tr>
<tr>
<td>RCCG10</td>
<td>EPI</td>
<td>Low latency/high performance services integrated.</td>
</tr>
</tbody>
</table>
Data Grids

Research challenges for:

- Computing grids
- Data grids
- Knowledge grids

- RCDG4: Improved distributed data management
- RCDG5: 
- RCDG7: 
- RCDG2: Distributed data models
- RCDG1: Quality of service
- RCDG3: 

The Share project (FP6-2005-IST-027694) is funded by the European Commission Information Society and Media DG under the 6th Framework Programme.
## Data Grids

<table>
<thead>
<tr>
<th>Challenge</th>
<th>Community</th>
<th>Description of the requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCDG1</td>
<td>EPI</td>
<td>• Easily installed grid middleware for health environments.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Low maintenance and administration.</td>
</tr>
<tr>
<td>RCDG2</td>
<td>EPI - VPH</td>
<td>• Data architectures/tools for private data dissociation, pseudo/anonymisation and encryption.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Automatic compliance with legal requirements.</td>
</tr>
<tr>
<td>RCDG3</td>
<td>EPI</td>
<td>• Exploitation models and guarantees QoS for services.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Advance resource reservation with pre-negotiated QoS.</td>
</tr>
<tr>
<td>RCDG4</td>
<td>EPI</td>
<td>• Scalable data cataloguing and data transfer.</td>
</tr>
<tr>
<td>RCDG5</td>
<td>VPH</td>
<td>• Storage services for easy upload/download of large binary objects.</td>
</tr>
<tr>
<td>RCDG6</td>
<td>VPH / EuroPhysiome</td>
<td>• Distributed data models and repositories multiscale data.</td>
</tr>
<tr>
<td>RCDG7</td>
<td>IMI</td>
<td>• Enhanced standards for data protection in web services environments.</td>
</tr>
</tbody>
</table>
Knowledge Grids

Research challenges for:

- Computing grids
- Data grids
- Knowledge grids

Grid technology challenges
Defining standards and ontologies
Research area challenges

The SHARE project (FP6-2005-1ST-027694) is funded by the European Commission Information Society and Media DG under the 6th Framework Programme.
## Knowledge Grids

<table>
<thead>
<tr>
<th>Challenge</th>
<th>Community</th>
<th>Description of the requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>RCKG1</td>
<td>EPI</td>
<td>Knowledge-driven grid catalogues and integration based on the metadata.</td>
</tr>
<tr>
<td>RCKG2</td>
<td>IMI</td>
<td>Standards and models to expose web services (semantics), scientific services, properties of data sources, data sets, scientific objects, and data elements</td>
</tr>
<tr>
<td>RCKG3</td>
<td>IMI</td>
<td>Enhanced knowledge representation models and data exchange standards for complex systems</td>
</tr>
<tr>
<td>RCKG4</td>
<td>IMI</td>
<td>Develop new, domain-specific ontologies based on standard data representation models and reference ontologies</td>
</tr>
<tr>
<td>RCKG5</td>
<td>IMI</td>
<td>Advanced text mining tools to capture implicit information about complex objects, relationships and processes, as described in patents and literature</td>
</tr>
<tr>
<td>RCKG6</td>
<td>IMI</td>
<td>Standards and an expert tool (ontology/schema/rules negotiator) to expose properties of local sources in a federated environment</td>
</tr>
<tr>
<td>RCKG7</td>
<td>IMI-VPH</td>
<td>Standards and an expert tool (services/data negotiator) to guide users through the complexities of the data, data models, simulation and modelling tools.</td>
</tr>
</tbody>
</table>
Collaboration Grid
For e-science/e-health

Data Grid
Distributed and optimized storage of large amounts of accessible data

Computing Grid
For data crunching applications

Revisiting User Reqs
Collaboration Grid
For e-science/e-health

Knowledge grids a level up

Data Grid
Distributed and optimized storage of large amounts of accessible data

Computing Grid
For data crunching applications

The SHARE project (FP6-2005-IST-027694) is funded by the European Commission Information Society and Media DG under the 6th Framework Programme.
Challenges & complexity

- Quality of Service
- User Friendliness
- On Demand Access
- Infrastructure Interoperability
- Improved Distributed Data Management
- Domain Specific Knowledge Management and Ontologies
- Knowledge Management Tools and Standards
- Data Integration Tools and Standards
- Distributed Data Models

Time vs. Complexity

The SHARE project (FP6-2005-IST-027694) is funded by the European Commission Information Society and Media DG under the 6th Framework Programme.

ISGC 2008 - Taipei - 7th - 11th April, 2008
Conclusion (1/2)

- Grid Technology has been identified as one of the key technologies to enable and support the "European Research Area"

- The impact of the Grid concept is expected to reach far beyond eScience, to eBusiness, eGovernment and eHealth

- Continuing and reinforced European and National R&D for HealthGrid services and for the deployment of dedicated grids infrastructures in the Biomedical & Healthcare world

- A major challenge is also to take the technology out of the Laboratory to the Citizen
Conclusion (2/2)

Through the EU funded EUAsiaGrid project we expect to:

- Extend this roadmap to the Asia-Pacific Region

- Assess similarities and differences between Asian-Pacific countries themselves

- Assess similarities and differences between Asia Pacific and Europe

Reminder:

- HealthGrid 2008 conference - June 2-4, 2008
Thank you for your attention!