



Enabling Grids for
E-science in Europe

www.eu-egee.org

NA4/SA1 WG – Kick-off meeting, Sept. 15th 2004

Aims and organization of the Biomedical VO

Yannick Legré
CNRS/IN2P3



EGEE is a project funded by the European Union under contract IST-2003-508833

Disclaimer

- These slides are a basis for discussion
 - will be subject to modification during this meeting (I look for input from you) ... and after!

Conclusion

of the AA meeting of September 13th.

Top five risks of the EGEE project (extracted from Fab's slides)

1. **Lack of buy-in to EGEE** by new sites / VOs, National networks and programmes, user communities and applications:
4. **Poor level of quality of service**
 - Low level of deployed nodes and sites
 - Expensive and complex VO management

(extracted from Cristina's slides)

- failure to attract new applications and user communities
- failure to integrate existing national and regional infrastructures
- failure to work with other international grid projects

(extracted from Bob's slides)

- **Still confusion about VO management/definition/deployment:**
 - Role/responsibility/scope of VO managers
 - VO creation/deployment process is not clear enough

What is a “VO”? (1/2)

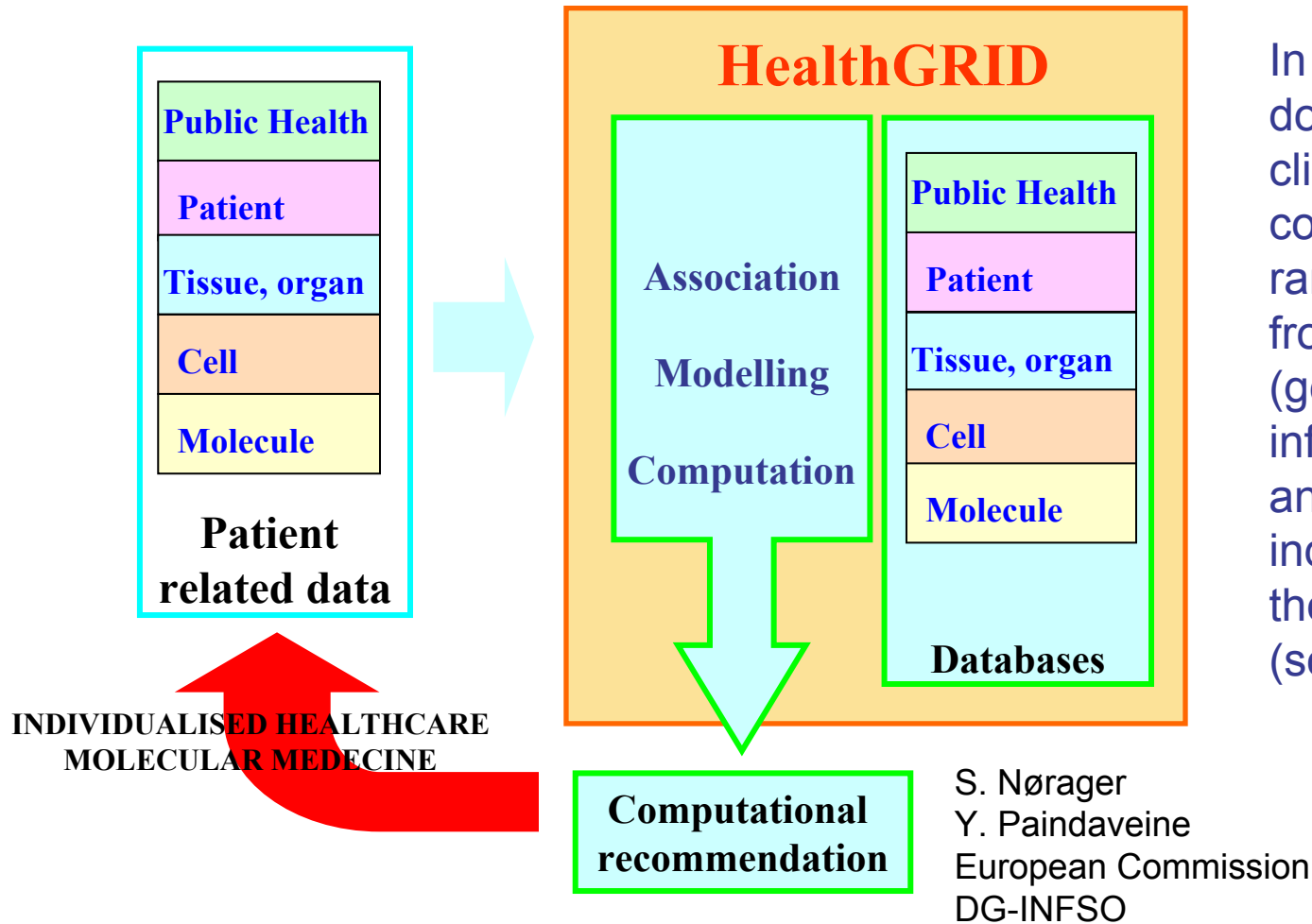
- A VO is a means to describe the organization of a community. This is a view which is not frozen and could evolve according to the modifications of the community it represent.
- There are approximately as many VOs concept as different user communities.

What is a “VO”? (2/2)

- A VO can be understood from different points of view: i.e.
 - Thematic:
 - Bioinformatics
 - Medical images
 - Medical informatics
 - Organisational/Geographical
 - Hospital
 - Collaboration
 - Region/Country

→ We need to agree on the definition of a VO and on the level of granularity we apply to this definition.

A look at the future: the HealthGrid vision



In this context "Health" does not involve only clinical practice but covers the whole range of information from molecular level (genetic and proteomic information) over cells and tissues, to the individual and finally the population level (social healthcare).

S. Nørager
Y. Paindaveine
European Commission
DG-INFSO

Aims of the Biomed VO

- To structure and organize the biomedical community which is:
 - Atomic (small groups of people)
 - More individualist
 - Focused on privacy and security issues
 - Not centralized organization (like CERN, ESA, ...)
 - Facing to new challenges
 - Not skilled in computer science
- Play the guinea pig to setup and deploy a non-HEP VO
- To put the solution into general use by setting up a new community integration procedure with SA1

Requirements

- To have a full set of services up and running as soon as they are available
- Fault tolerant infrastructure (i.e.: duplication of core grid services)
- User and admin support
 - The first level of user support can probably be done with relevant and more explicit error messages...
- Automatic tool to install/configure grid middleware inside application institute
 - Phase 2 – when the new VO will move in production phase and will compromise some resources into the Grid.
- Quality metrics
- ...

Deployment of a VO

- To be fully functional a VO needs to have:
 - RB + BDII (*)
 - RLS (*)
 - CE + WN
 - SE
 - VO LDAP server (*)
 - VO manager + deputy
 - Access to a MyProxy server (useful for long jobs, not mandatory)
- To increase fault tolerance and avoid bottleneck, the Core Grid Services (*) should be, at least, duplicated.

- Who is in charge to setup core grid services for a VO ?
 - CIC as a whole ?
 - Does the VO negotiate with the different CICs or shall we transmit the request to a representative?
- Needs to define a workflow channel to inform users/biomed loose canons about:
 - Maintenance operation
 - Adding of new resources or services
- Needs to setup support for VOs
 - User support
 - Admin support

Issues (2/2)

- Define % resources from EGEE / resources compromised by the VO
- Define policy for usage of EGEE for new comers.
- How biomed loose canons will interact with SA1 staff?
- ...

Quality issues

- **Quality metrics allow to quantify the growth of EGEE user communities and their experience of the infrastructure**
- **Metrics described in NA4 quality plan: target values from Technical Annex**
 - Extension and growth of user communities: number of disciplines, number of users, number of “scientific” VOs, geographic extension of the VOs
 - Experience of the infrastructure: quality of service relative to job success rate and execution time, VO specific usage of the infrastructure,...
 - Network usage, user support, test phases, ...
- **Requirements added (M. Soberman) to data base regarding**
 - R0047: NA4 Metrics and QoS relative to jobs requests and execution
 - R0049: NA4 Metrics requirement - Abort codes classification and recording
 - R0050: NA4 Metrics: Network usage