CompChem VO progress report

NA4 Generic Applications Meeting – Jan 9th, 2006 Catania, Italy

Osvaldo Gervasi[1], Antonio Laganà[2],

[1] Dept. of Math. & Computer Science, University of Perugia

[2] Dept. of Chemistry, University of Perugia

Content

- Overview of CompChem VO
- CompChem VO Management
- GEMS
- Basic requirements
- The Computational Chemistry Community
- The VO evolution
- Current status and future developments

Overview of CompChem VO

- The Virtual Organization CompChem supports the needs of the Computational Chemistry community in the EGEE Grid.
- The VO started on September 2004
- The VO originates from the experience made in the COST in Chemistry D23 Action METACHEM: Metalaboratories for complex computational applications in chemistry, in which cooperative research and knowledge handling in chemistry were promoted through the funding of Collaborative Laboratories, called Metalaboratories
- The GEMS application, the core project of the VO, was deployed as a demo for the EGEE Reviewers in January 2005

CompChem VO

- **A Memorandum of Undestanding has been established between EGEE and CompChem**
- **A FP6 project has been submitted and positively evaluated**
- **The main targets of the VO are:**
  - Promote molecular simulations
  - Consolidate GEMS:
    - the calculation and fitting of the electronic energies of atomic and molecular aggregates using high level ab initio methods
    - the use of statistical kinetics and dynamics means to study chemical processes
  - Extend the membership to other competencies and laboratories
  - Develop the handling and creation of molecular and chemical e-knowledge for research and education.
  - Design an evaluation model of credits and Grid service costs to reward partner laboratories

CompChem VO management

- **The management of the VO is based on the collaboration between two Departments of Perugia University**
- **Chair: Prof. Antonio Laganà, Dept. of Chemistry**
- **VO manager: Dr. Osvaldo Gervasi, Dept of Maths and Computer Science**
- **Support team:**
  - Dr. Cristian Dittamo
  - Dr. Matteo Diarena
  - Dr. Francesca Gentili

CompSci

  - Dr. Leonardo Pacifici
  - Dr. Leonardo Arteconi
  - Dr. Federico Filomia

Chemistry

CompChem VO management

- **The VO is based on the sharing of some basic assets and on the active involvement of the partner Laboratories**
- **A Consortium agreement has to be signed by each partner**
- **Each partner may be involved at different levels:**
  - **User:** implementation on the Grid of a suite of codes of exclusive interest for the implementing laboratory
  - Code offer: the laboratory confers to the VO a stable suite of codes
  - Service offer: the laboratory participates to the management of the Grid infrastucture (manpower, hardware, service brokering and monitoring, etc), the development of joint projects etc.

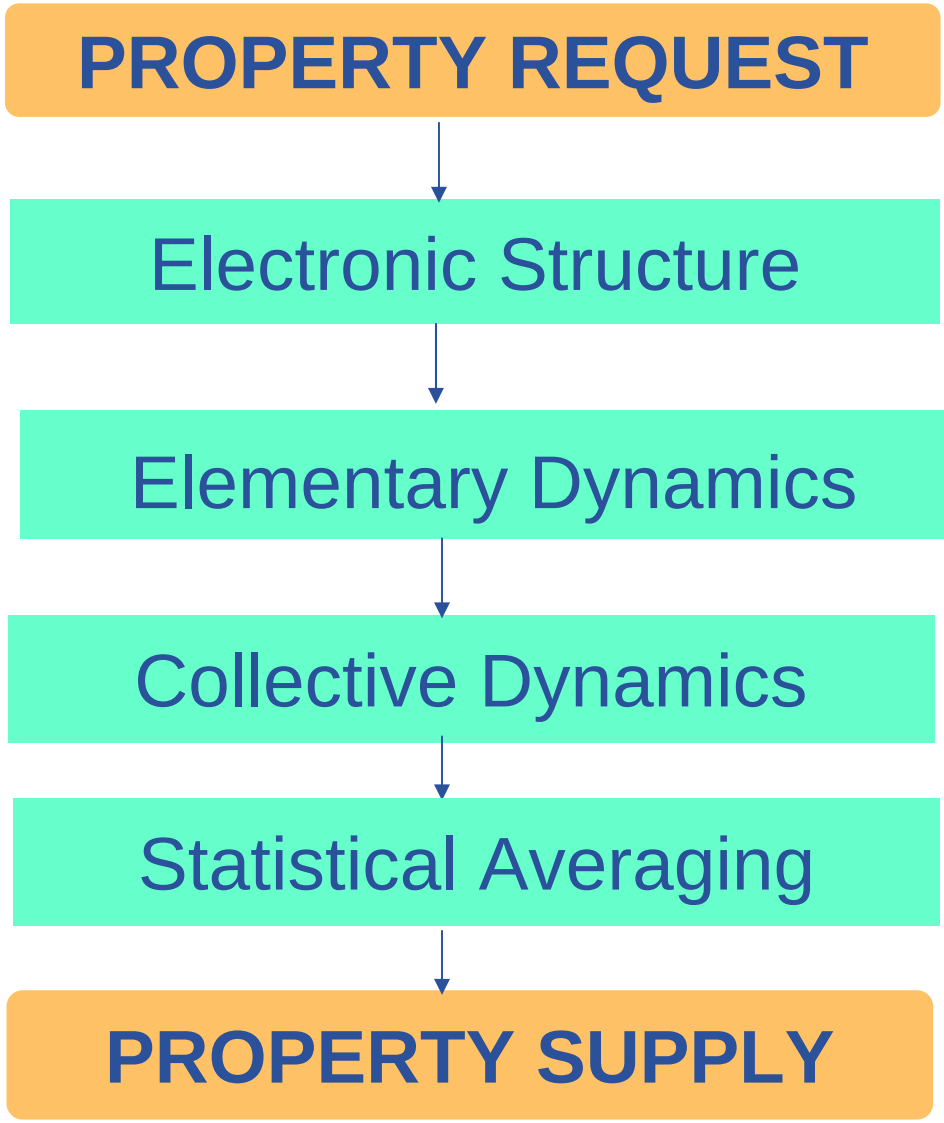
The Grid Enabled Simulator (GEMS)

- **Objective: Calculation of the properties of the molecular processes using a priori computational approaches**
- **GEMS.0: deals with atom diatom reactions and calculation of properties using quasiclassical means**
- **GEMS.1: moves to a more synergistic application model, by:**
  - Including ab initio calculations and fitting of the potential values (when needed)
  - Increasing the number of atoms that can be dealt with classical methods
  - Increasing the number of participant labs and the related know how

GEMS project

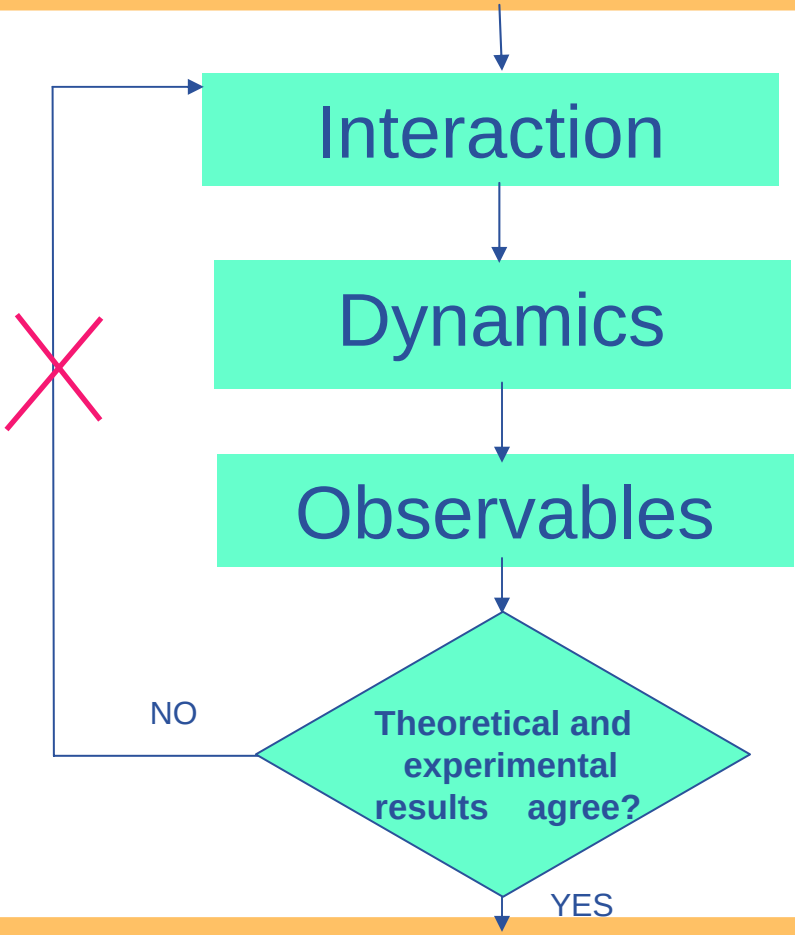
- **GEMS.2:**
  - Include quantum and semi-classical approaches to dynamics calculations
  - Include new approaches and packages
  - Increase the number of virtual monitors
  - Increase the Molecular Virtual Reality components and use of immersive tools
  - Articulate the packages management
  - Further increase the number of participant laboratories and the related know how

GEMS architecture



The GEMS.0 demo application

REQUEST: a potential fitted to beam experiments



SUPPLY: the potential and related monitors

GEMS.0 implementation on GILDA

end 2004

- **SPECIFIC ASSUMPTIONS OF THE GEMS.0 DEMO**
  - Focus on atom diatom systems
  - Take potential energy routine from a library
  - Adopt classical dynamics (trajectories)
  - Runs on the GILDA testbed infrastructure
- **GEMS.0 EXECUTION ON GILDA**
  - Porting on GENIUS Portal (Emidio Giorgio, INFN, Catania)
    - Common layout of the other applications
    - Key services
      - *Resource Broker:* *grid004.ct.infn.it*
      - *Computing Element:* *ce.grid.unipg.it*
      - *User Interface:* *grid-tutor.ct.infn.it*
      - *GENIUS Portal:* *https://genius.ct.infn.it*
  - Web portal on the local User Interface (Cristian Dittamo, Matteo Diarena, Francesca Gentili, University of Perugia)
    - Key services
      - *Resource Broker:* *grid004.ct.infn.it*
      - *Computing Element:* *ce.grid.unipg.it*
      - *User Interface:* *ui.grid.unipg.it*

GEMS.1 implementation

may 2005

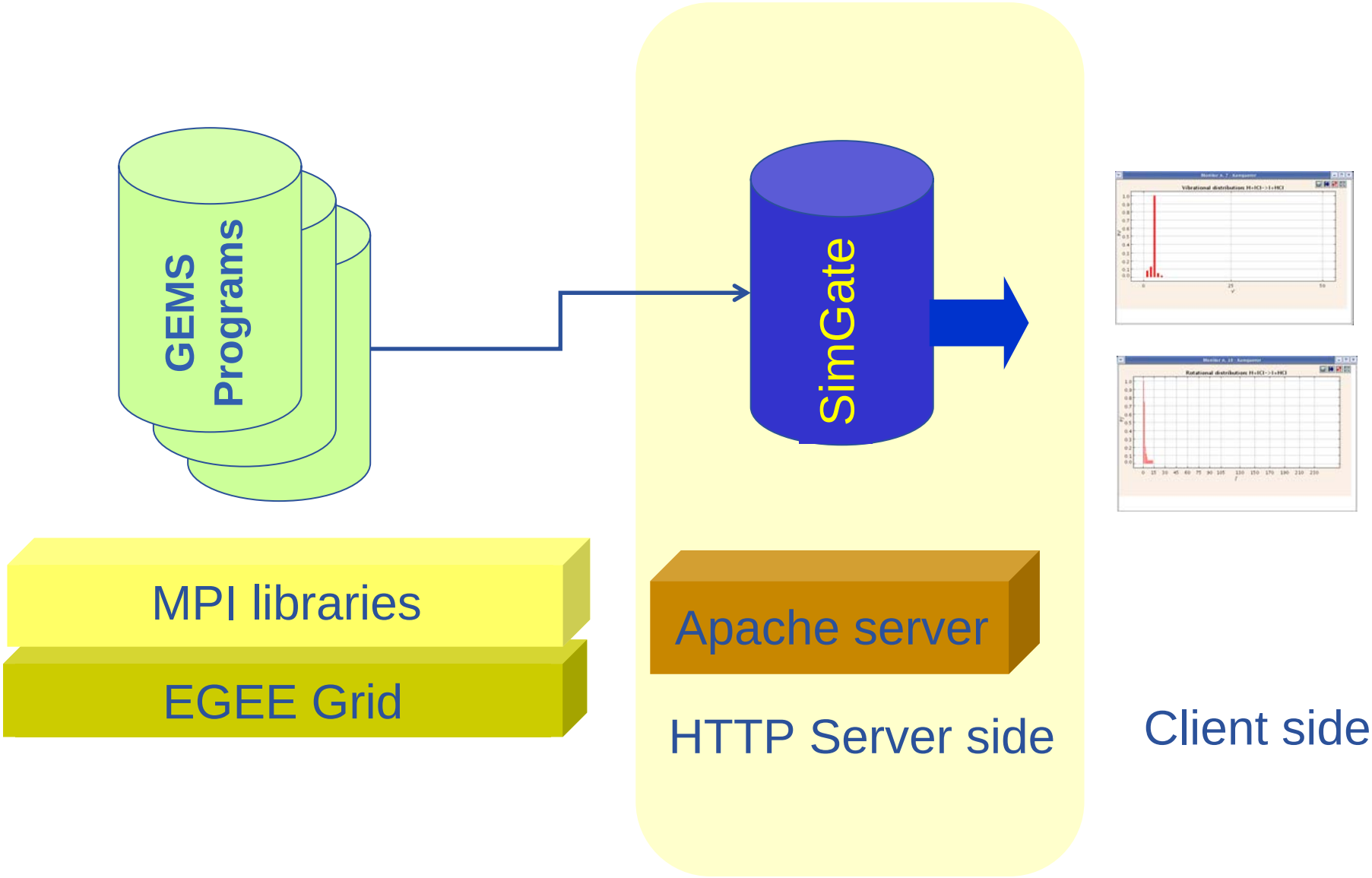
- **UNI-Perugia site included in the EGEE production environment**
  - Cluster of 14 nodes (biproc. PIII, 2GB RAM + 40GB HD per node) + CE + SE
  - Central management of Cluster services
  - Development of a Web site interfaced with the UI
- **Site management and administration**
- **User support**
- **Upgrade to LCG 2.6.0**
- **Inclusion of the Venus package for dealing with polyatomic systems**
- **Implementation of a prototype version of Potential Energy values calculations (SUPSim) and fitting.**

GEMS.2 implementation

end 2005

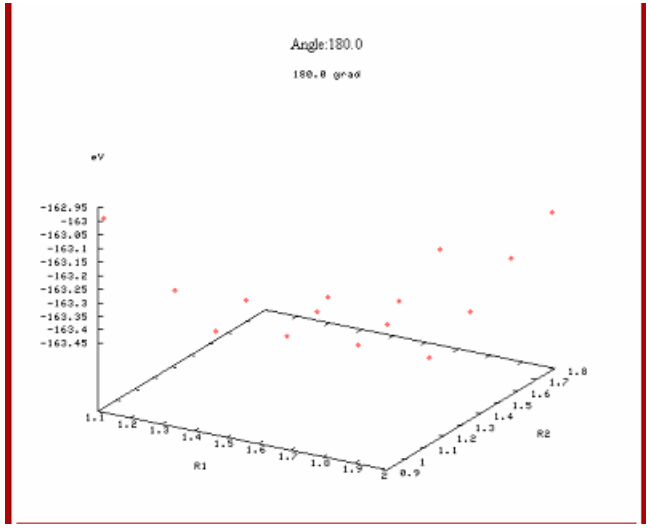
- Planning the inclusion of new resources in the production Grid (University of the Basque Country, Spain)
- Sites management and administration
- User support
- Inclusion of DL-POLY package for dealing with complex systems
- Implementation of RWAVEP code for Quantum Time Dependent approach
- Designing the implementation of the Columbus suite (Prof. Hans Lischka, University of Vienna)

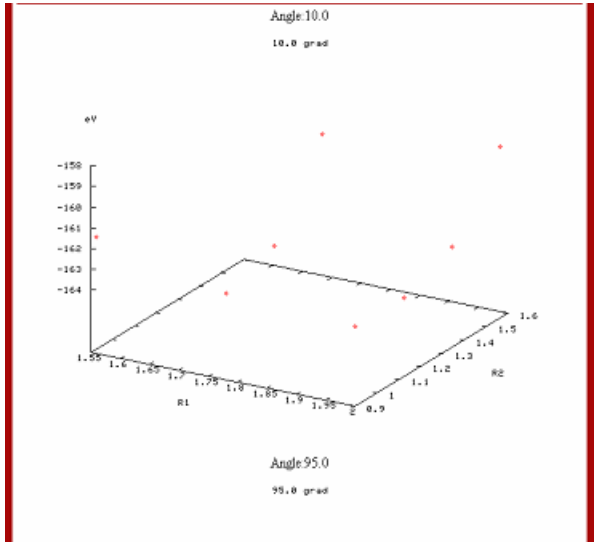
The model architecture of GEMS



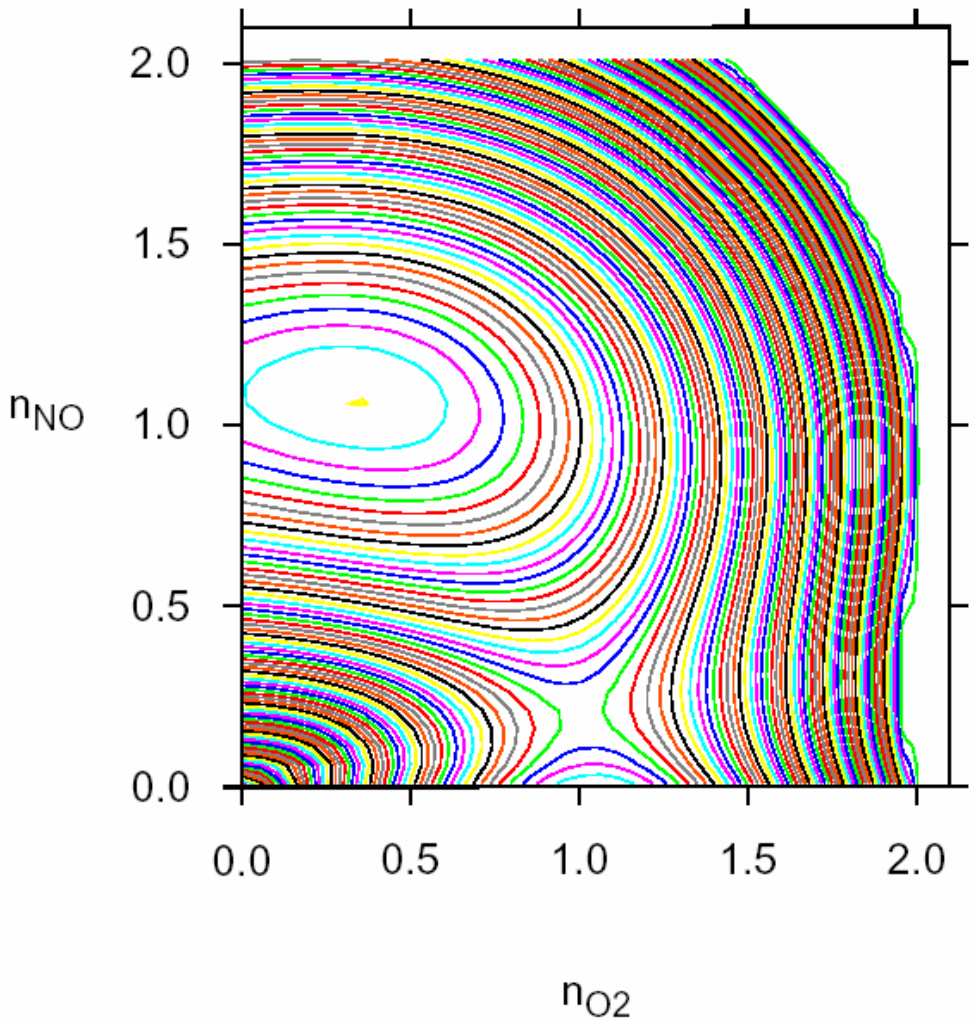
SupSim output:

View Results [N3 try script]	
Atom A1	N
Atom A2	N
Atom A3	N
Basis set	cc-pvdz
Method	Valence CASSCF
More details	
Spin [N]	
Spin [N]	
Spin [NN]	
Spin [NN]	
Spin [NNN]	4
Starting angle	10.0
Final angle	180.0
r1 Min	0.9
r1 Max	2.0
r2 Min	0.9
r2 Max	2.0
PES symmetry	A"
PES	

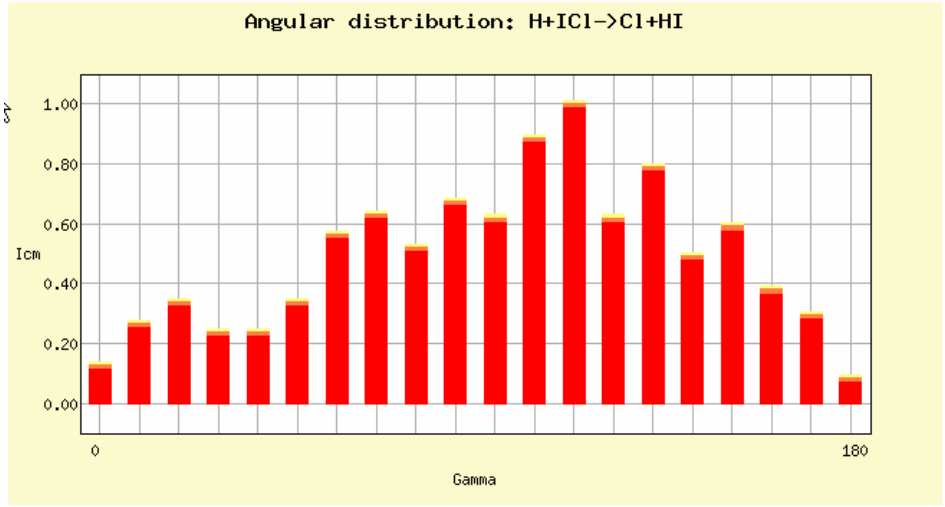


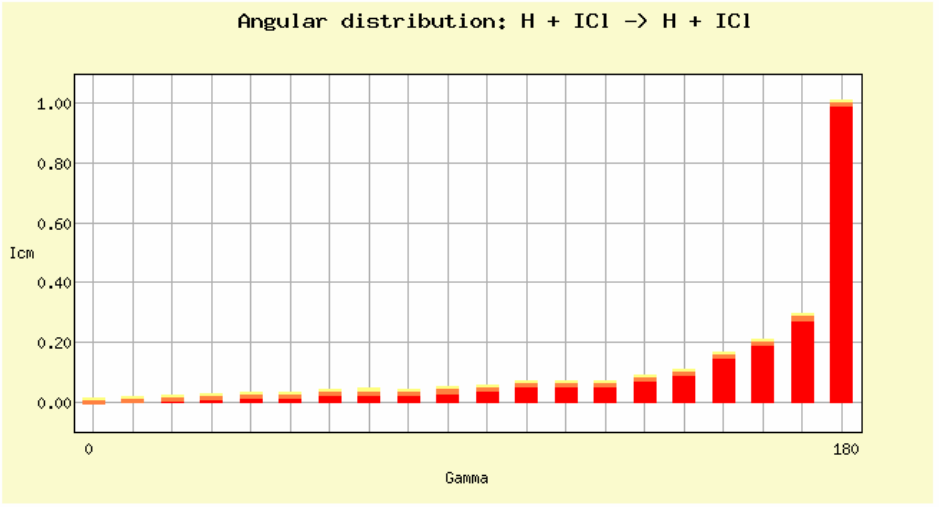


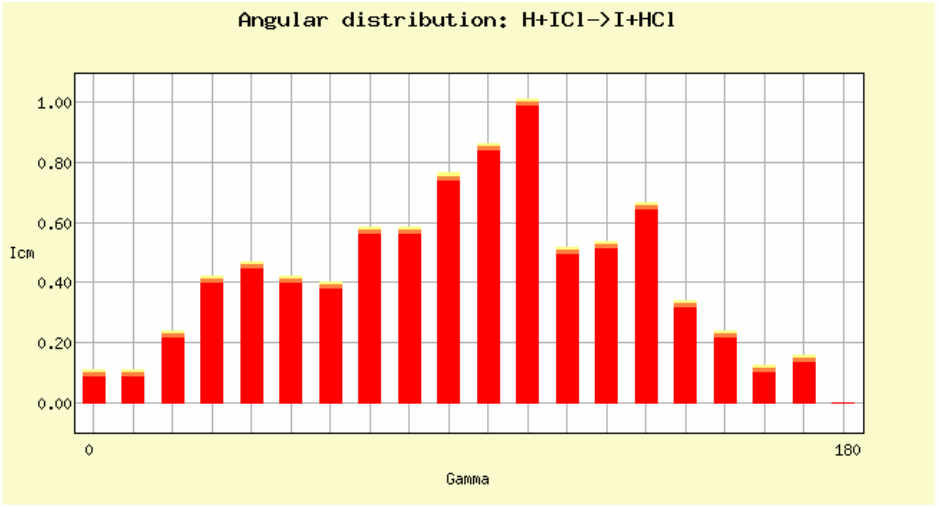
Fitting: $N+O_2$



The VIRTUAL MONITORS:
ANGULAR DISTRIBUTIONS

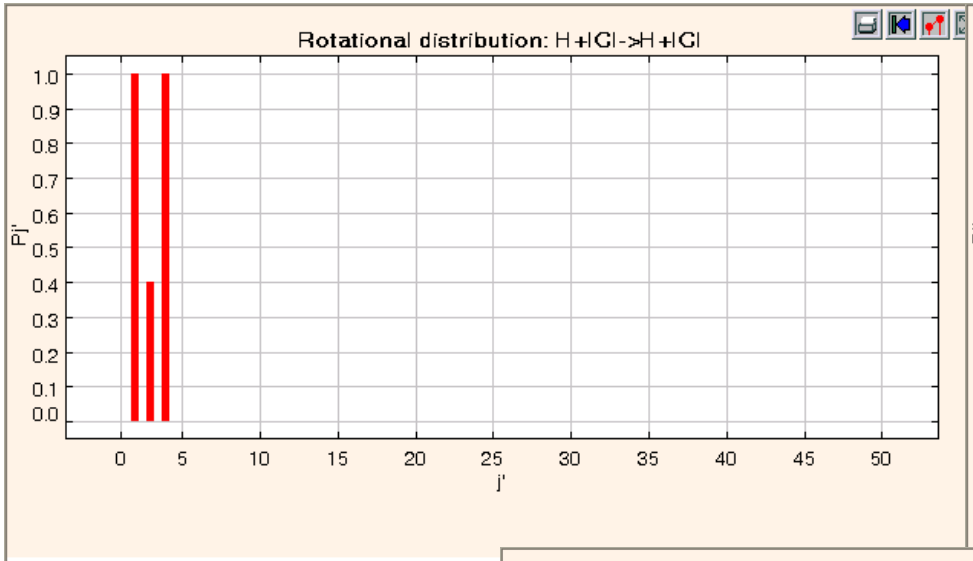


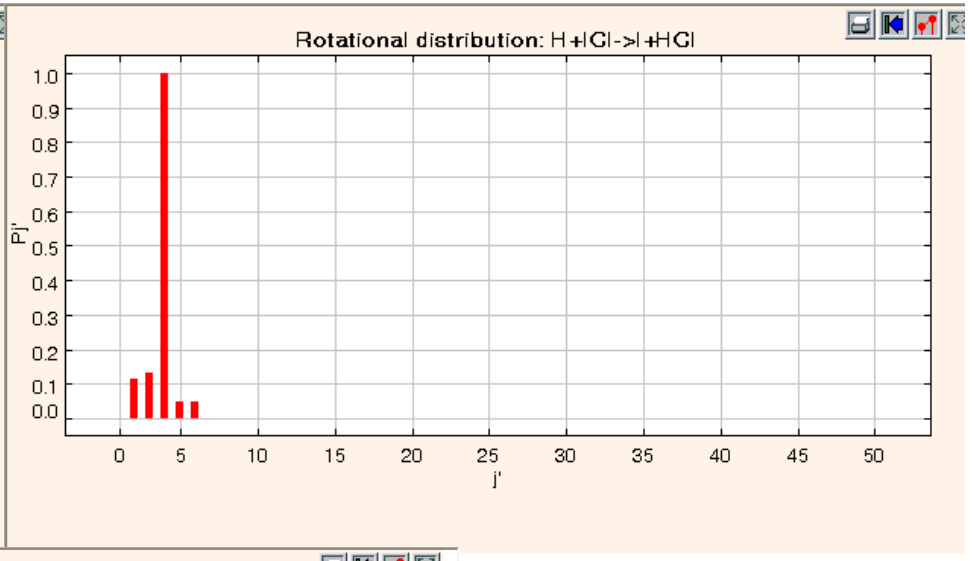


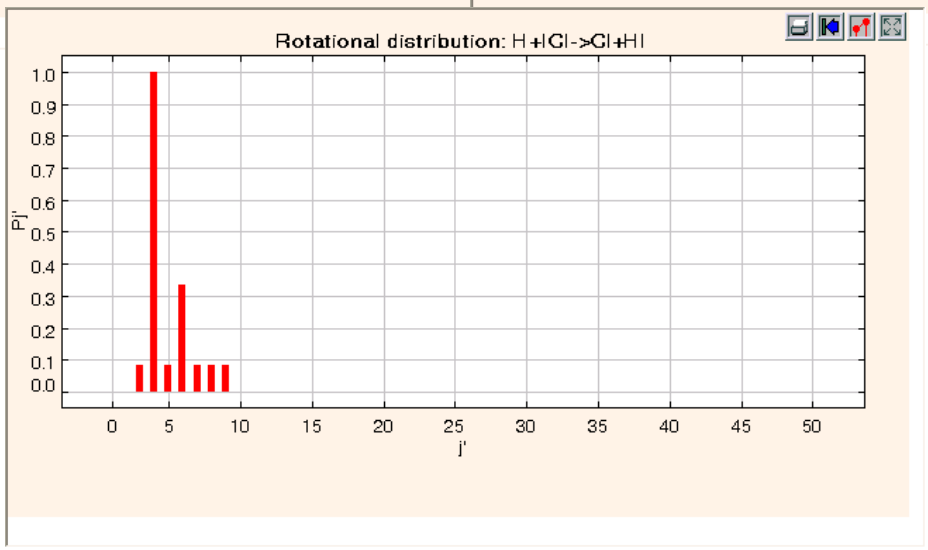


The VIRTUAL MONITORS:

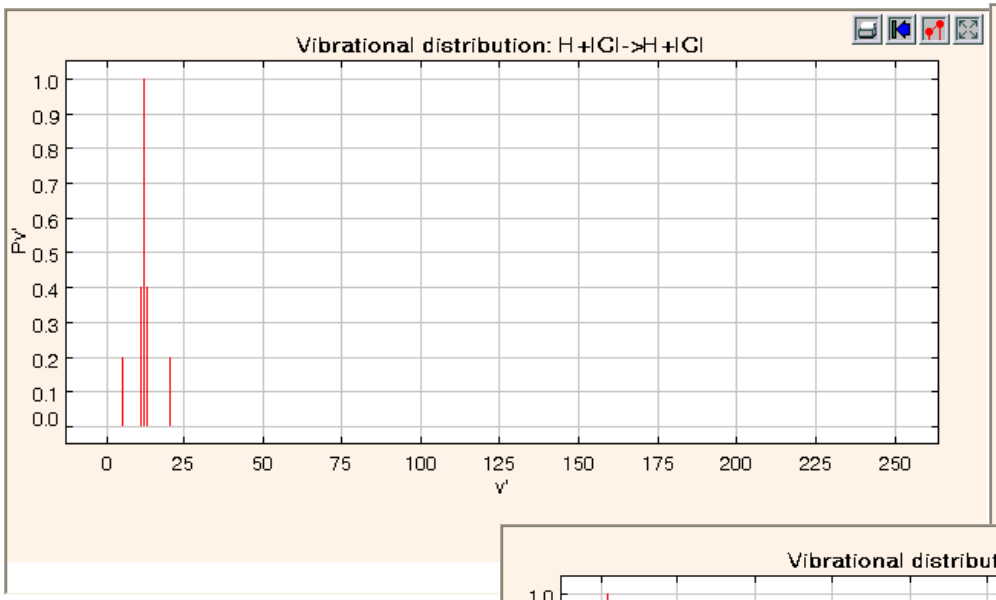
VIBRATIONAL DISTRIBUTIONS

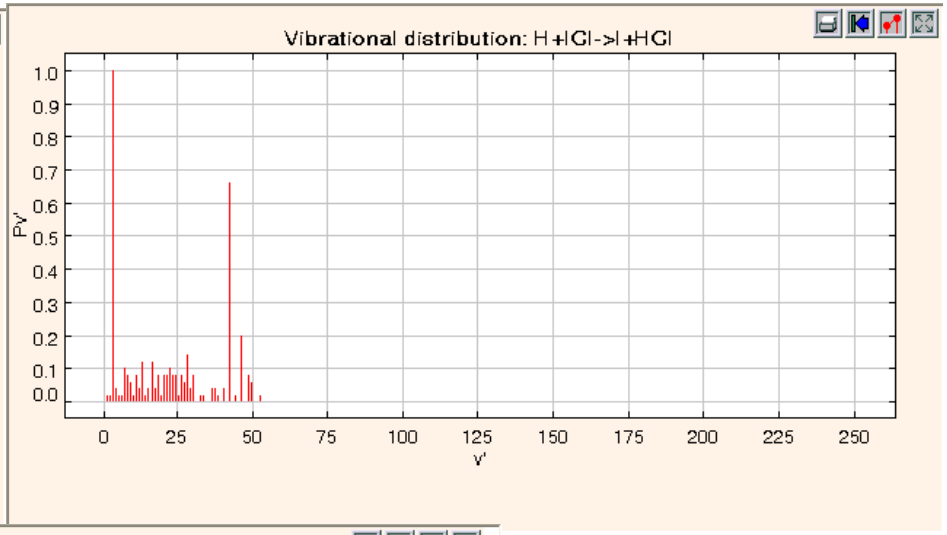


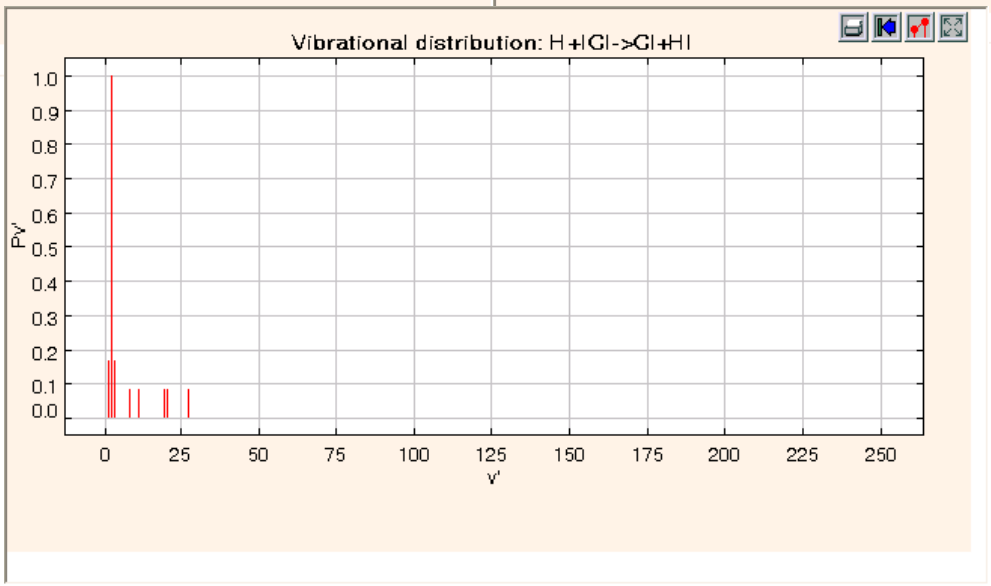




The VIRTUAL MONITORS:
ROTATIONAL DISTRIBUTIONS

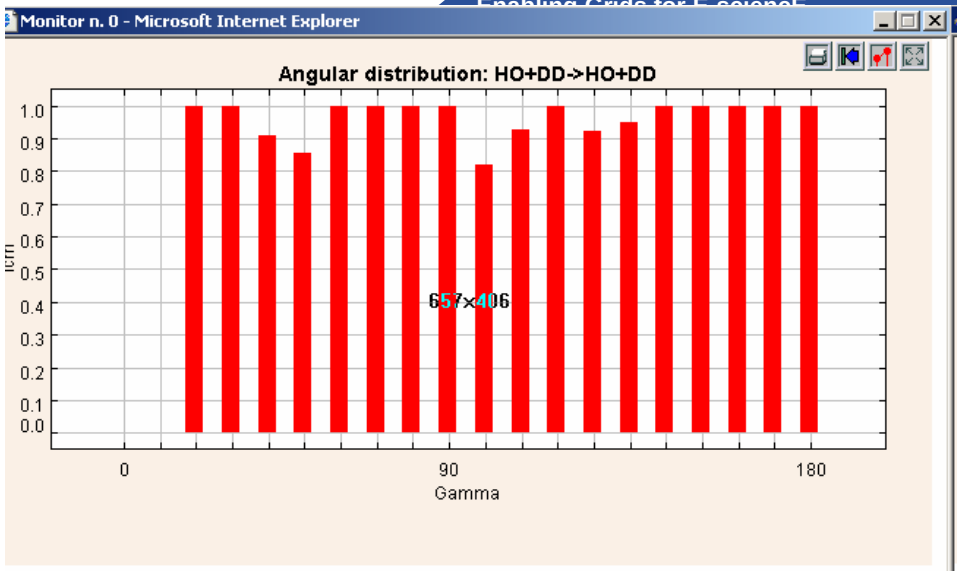


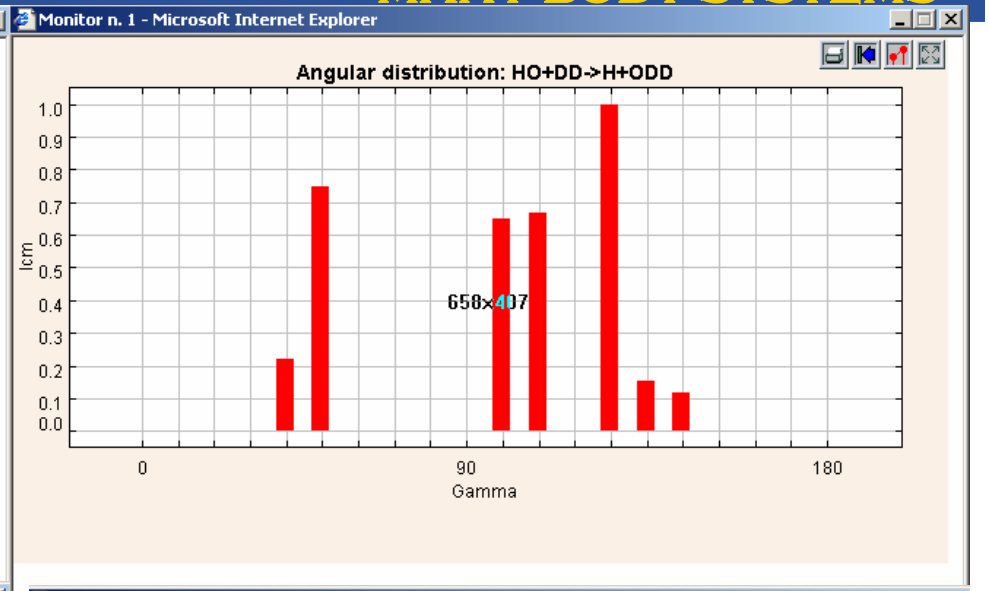


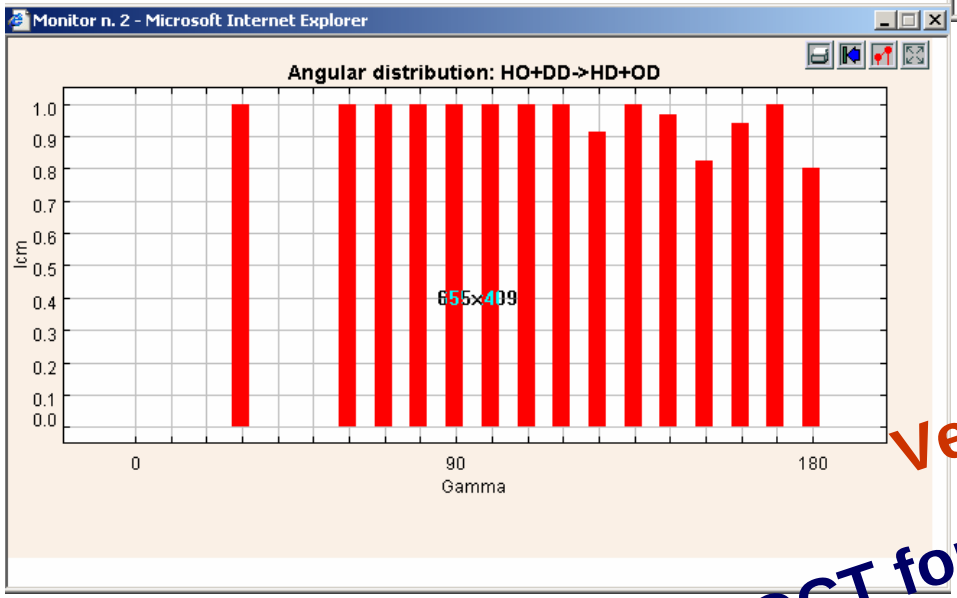


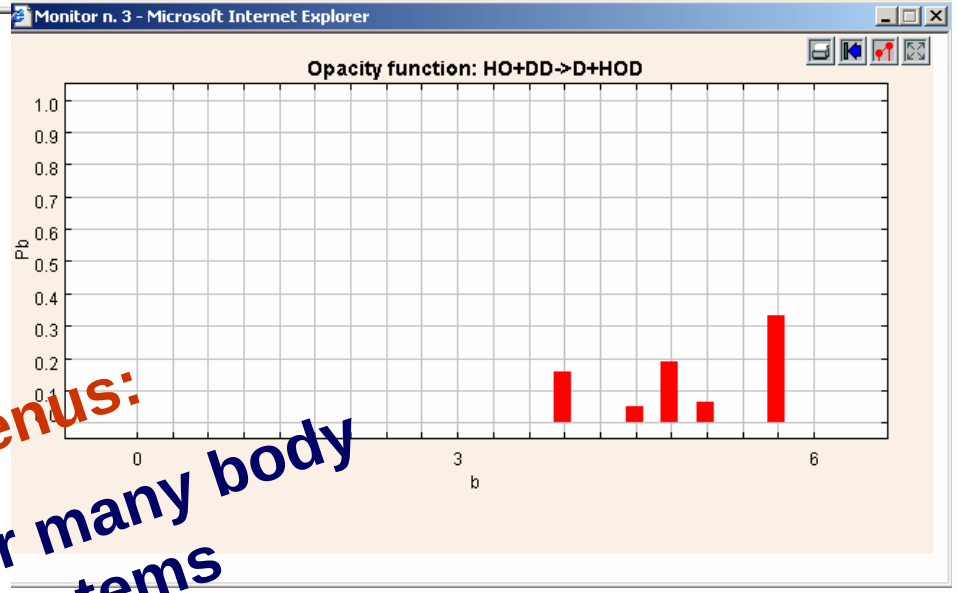
The VIRTUAL MONITORS:

MANY BODY SYSTEMS









Venus:

QCT for many body systems

OH-HCl trajectory

QCT for many body systems

Technical requirements

- We are mainly interested in using groups of clusters
- MPI must be available on the nodes to execute the various component packages in parallel
- MPICH-G2 should be available to allow the scheduling among the nodes of different CEs
- Scheduling should clearly privilege requests for a large number of nodes
- Licensed software must be requested via JDL and declared by the software enabled nodes
- The access to working nodes is required in order to guarantee the communication with the portal

The Grid added value

- **Software integration into distributed workflows**
  - to assemble applications out of various (different or complementary) distributed competences coordinated via the Grid (in the prototype electronic structure, elementary dynamics, statistical averaging, interfacing the experiment competences are involved)
- **Computational Campaigns**
  - to evaluate properties depending on the fate of few out of millions, billions or even more events by distributing the execution of the computations on the Grid (in the prototype these are some reactive events having specific energetic or steric properties)
- **Collaborative Engineering of knowledge**
  - to handle chemical information and knowledge including training and production of new knowledge
- **Security Infrastructure**
  - The state-of-the-art tools to share computational resources and to share computational codes among institutions in a secure fashion.

The Molecular Science Community of the MoU

Organization	Actor	Contact email
University of Barcelona	Margarita Alberti	maw@qf.ub.es
University of the Basque Country	Ernesto Garcia	qfpgapae@vc.ehu.es
Hungarian Academy of Sciences	Gyorgy Lendvay	lendvay@chemres.hu
University of Geneva	Laura Gagliardi	laura.gagliardi@unipa.it
IMIP- Italian National Research Council	Mario Capitelli	m.capitelli@area.ba.cnr.it
University of Goteborg	Gunnar Nyman	nyman@phc.gu.se
University of Crete	Stavros C Farantos	farantos@iesl.forth.gr
University of Vienna	Hans Lischka	hans.lischka@univie.ac.at
Swiss Federal Institute of Technology (ETH)	Hans Peter Luethi	hans.luethi@sl.ethz.ch

Planning for resources of the MoU

Resource nature	Estimated need till the end of EGEE	Resources committed by CompChem	Additional resources requested from EGEE
Storage (GEMS.1)	100KB / node	1.5 GB (15 nodes)	1.5 GB (15 nodes)
Storage (GEMS.2)	50GB / node	500 GB (10 nodes)	500 GB (10 nodes)
RAM on nodes (GEMS.1)	1MB / node	150MB (15 nodes)	150MB (15 nodes)
RAM on nodes (GEMS.2)	2GB / node	20GB (10 nodes)	20GB (10 nodes)
CPU time (GEMS.1)	1 week/month /node	15weeks/month (15 nodes)	15weeks/month (15 nodes)
CPU time (GEMS.2)	2 weeks/month /node	20 weeks/month (10 nodes)	20 weeks/month (10 nodes)

Deployment layout

Date	Milestone	Task
		Gilda Deployment
		Application development
Sept 30, 2005	GEMS.1 deployment on the production grid	Deployment on EGEE
Nov 30, 2005	GEMS.1 tested	Test
Jan 31, 2006	GEMS.1 on production	Production
Jan 31, 2006	GEMS.2 development	Application development
Mar 30, 2006	Metrics of satisf. of GEMS.1	Metrics of satisfaction
Apr 30, 2006	GEMS.2 deployment on the production grid	Deployment on EGEE
Jul 31,2006	GEMS.2 tested	Test
Oct 31, 2006	Metrics of satisf. of GEMS.2	Metrics of satisfaction

Current status and future developments

- **The difficulties of operating under the unfunded status**
- **Heavy manpower demand for routinary Grid operations**
- **High request for supporting end-user applications**
  - Nanotubes
  - Life sciences
  - Statistical Thermodynamics
  - Molecular Virtual Reality
- **Standard environments and training to get started**
- **Sustainability of the cooperation**
- **gLite migration has been just started**