



Enabling Grids for E-science

GEMS, Grid Enabled Molecular Simulations

EGEE 1st EU Review – 9th to 11th February 2005

CERN

Antonio Laganà¹, Osvaldo Gervasi²

¹ Dept. of Chemistry, University of Perugia

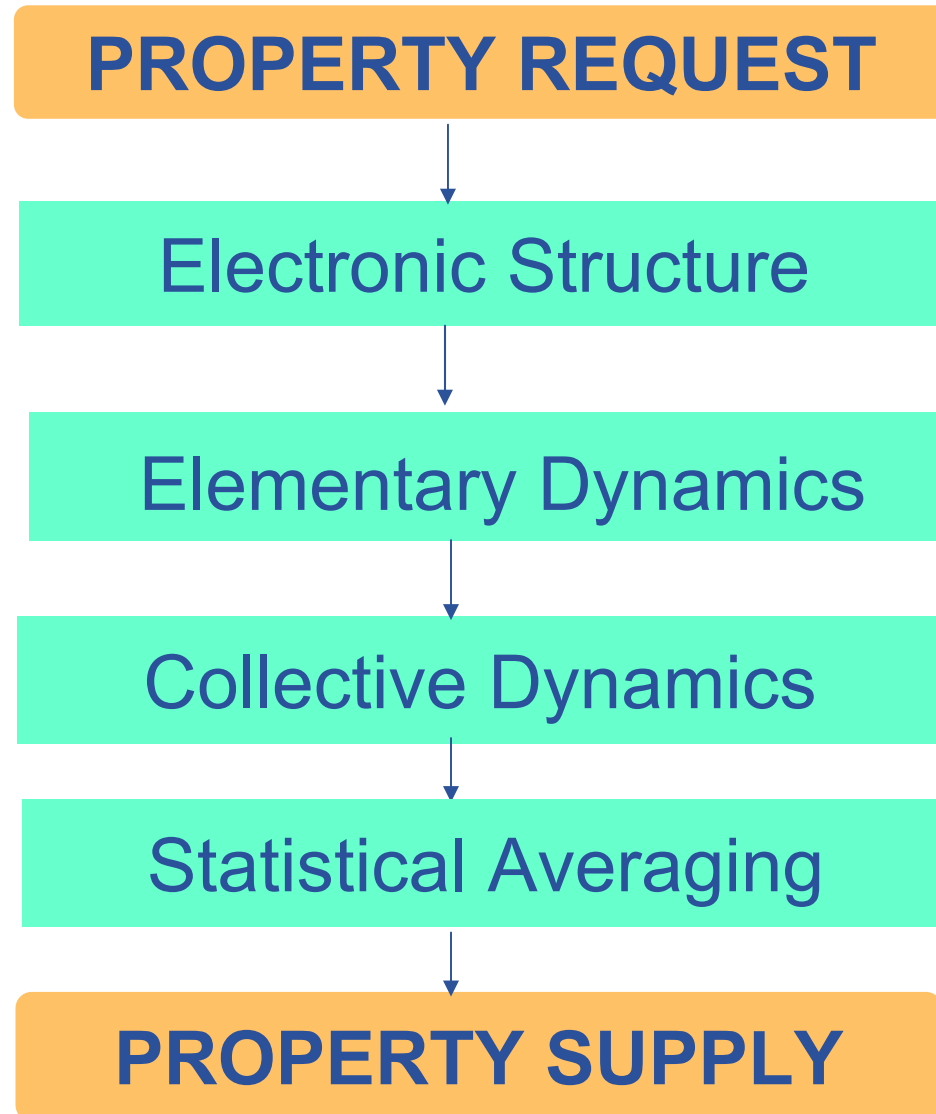
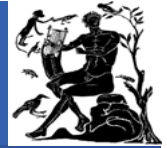
² Dept. of Math. & Computer Science, University of Perugia

www.eu-egee.org



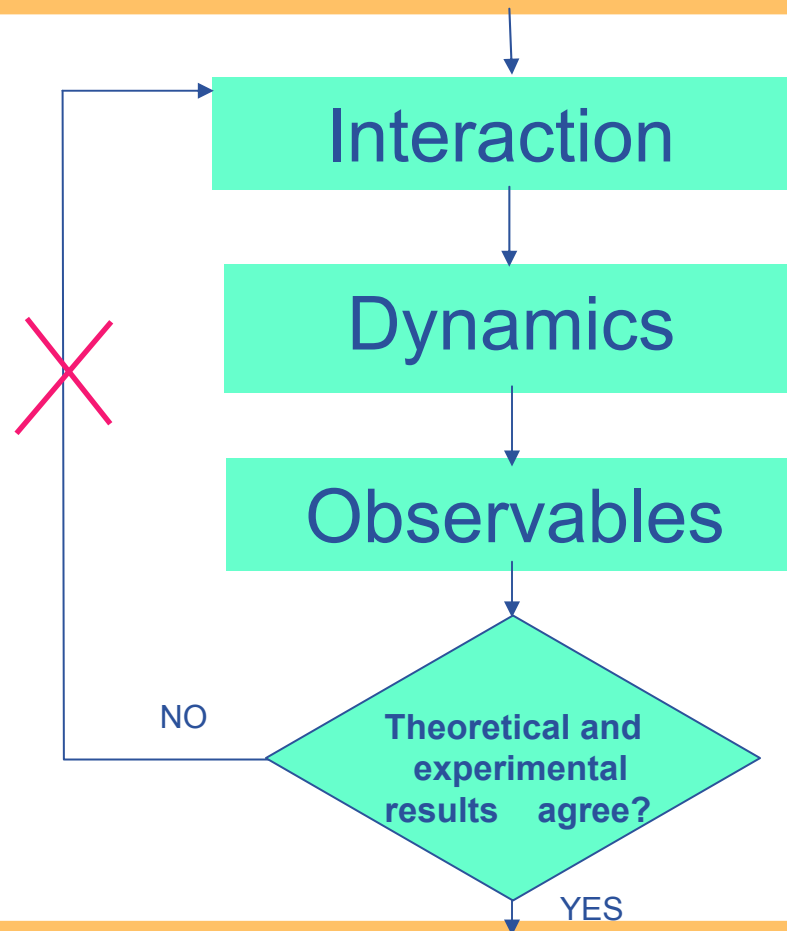
Università di Perugia

- **Overview of GEMS and the prototype application**
- **Deployment layout**
- **Scale of usage**
- **Usage of LCG-2 middleware services**
- **Difficulties and workarounds**
- **Cost comparison with alternative solutions**
- **Future work and direction**
- **Challenges for full production**
- **The Chemistry and Molecular Science Community**
- **The grid added value**





REQUEST: a potential fitted to beam experiments



SUPPLY: the potential and related monitors

- **SPECIFIC ASSUMPTIONS OF THE GEMS.0 DEMO**
 - Focus on atom diatom systems
 - Take potential energy routine from a library
 - Adopt classical dynamics (trajectories)
- **GEMS.0 EXECUTION ON GILDA**
 - Set the demo environment
 - Input the atoms involved
 - Input energy, initial vibrational and rotational state
 - Loop on trajectories by generating randomly the other initial conditions
 - Collect results as virtual monitors (in the real application to be compared with the experiment)



$$i\hbar \frac{\partial}{\partial t} \Psi(\{W\}, \{w\}, t) = \hat{H} \Psi(\{W\}, \{w\}, t)$$

Separation of electronic and nuclear motions

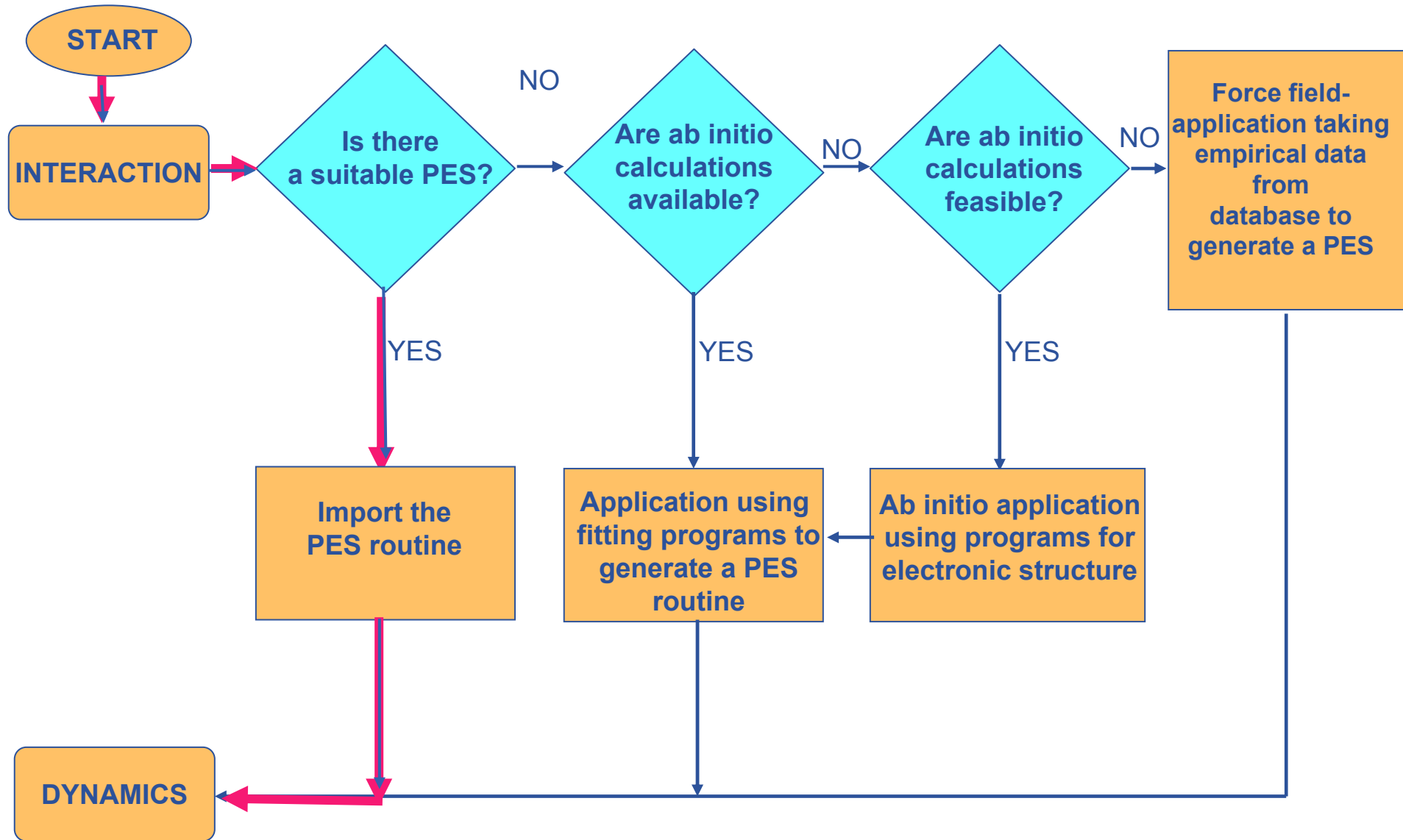


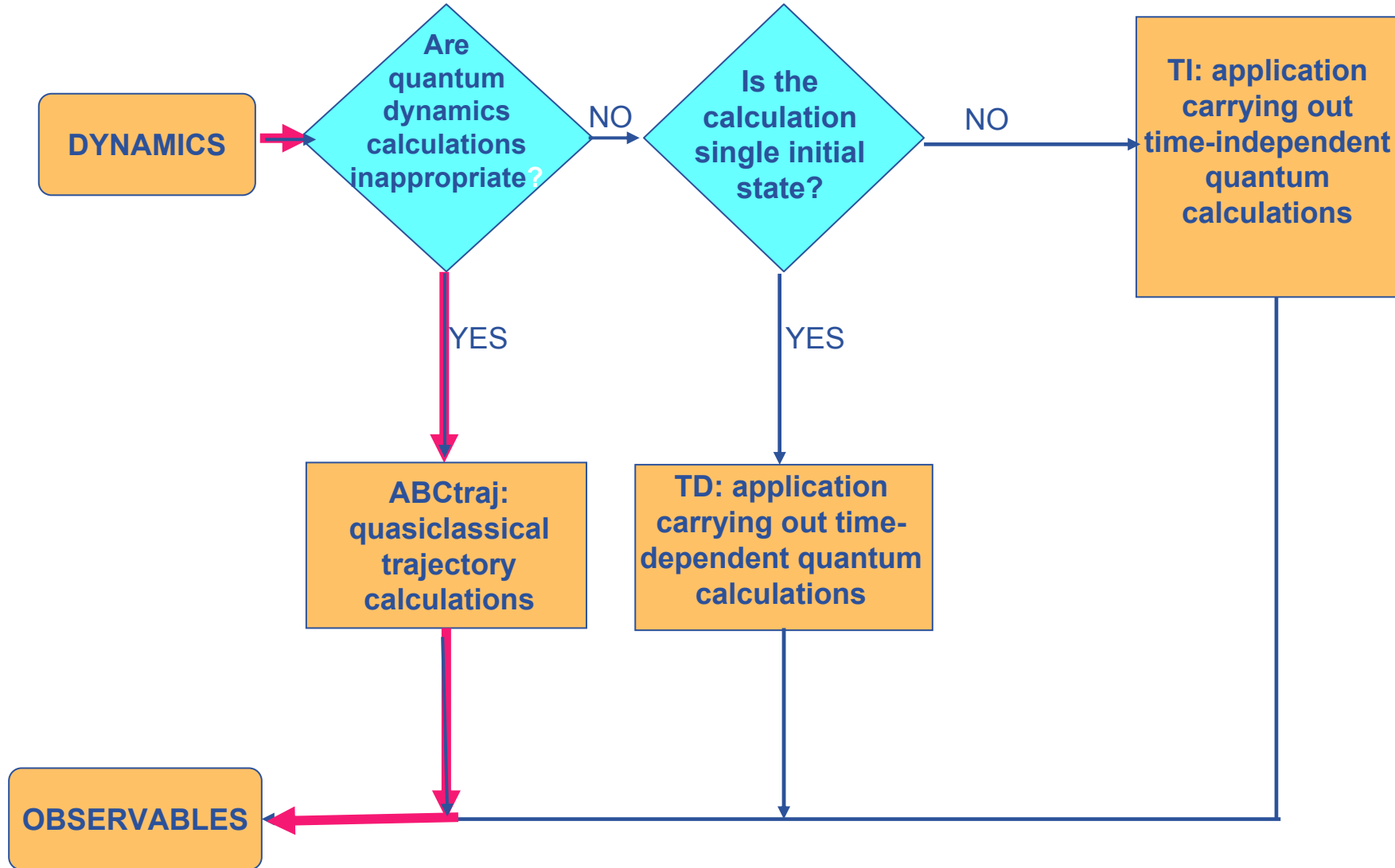
Electronic Schrödinger equation:

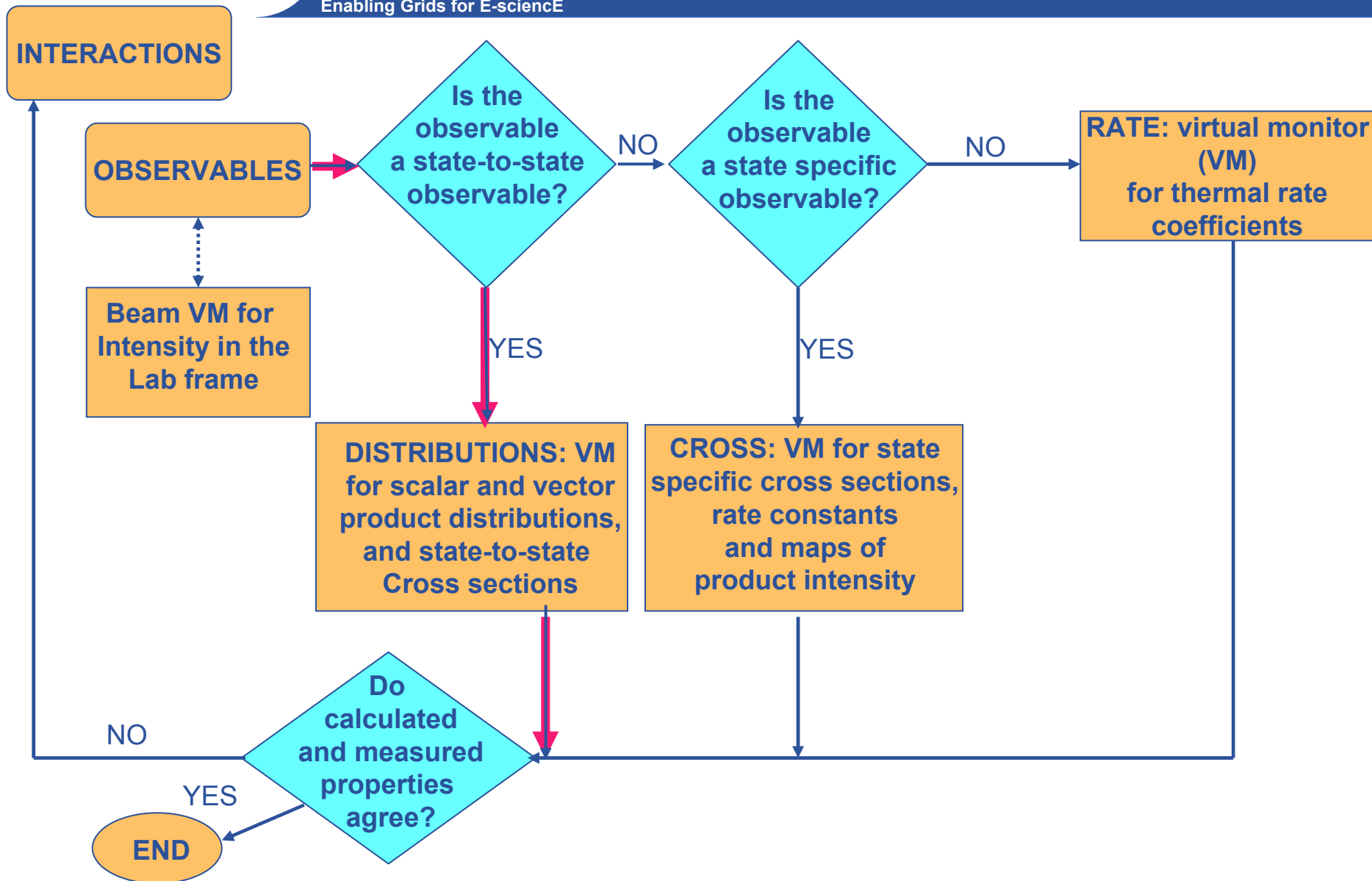
Nuclear Schrödinger equation:

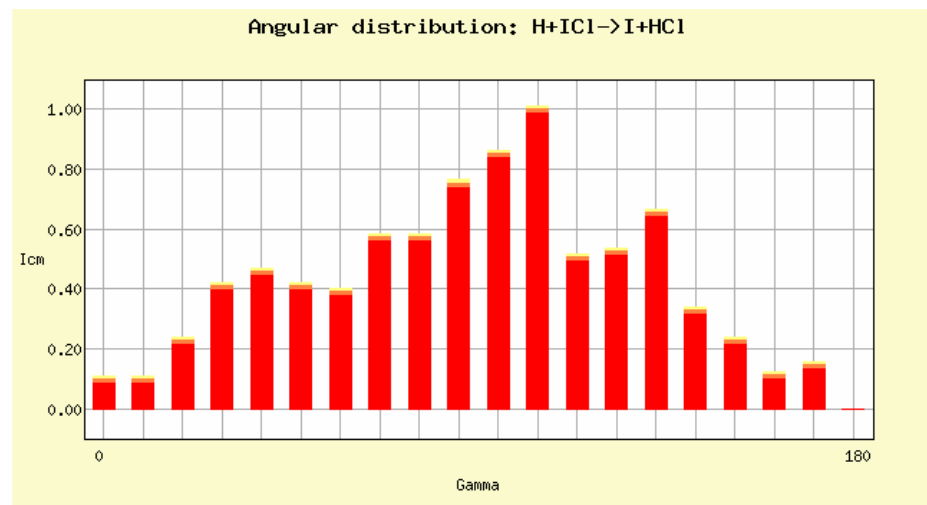
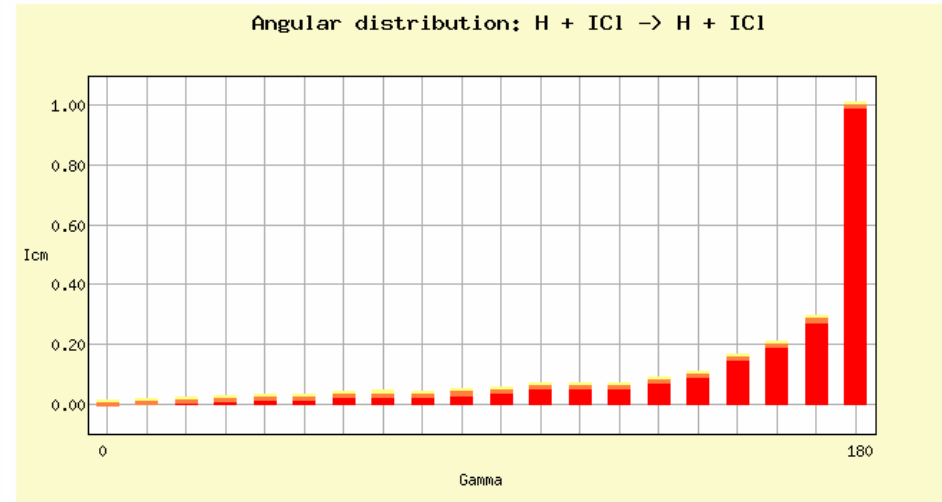
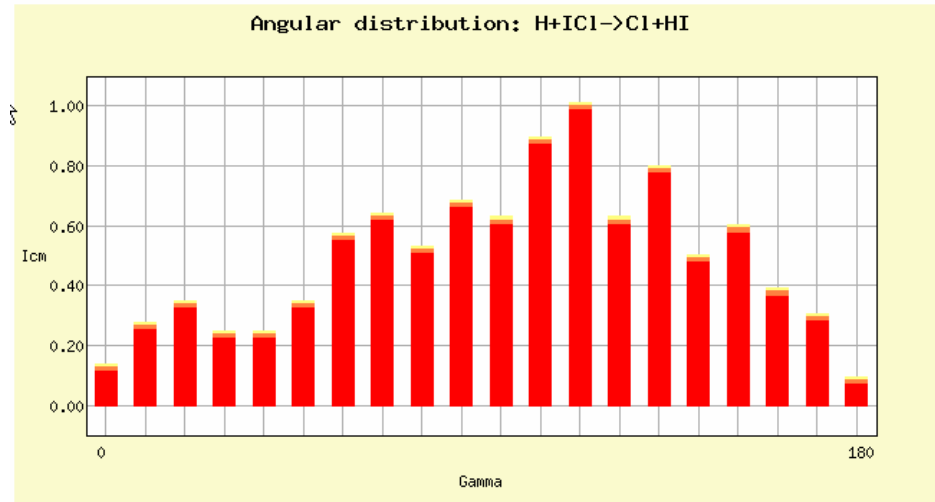
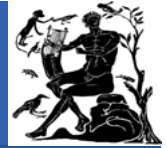
$$\hat{H}_{elec} \Psi_n(\{w\}; \{W\}) = E_n(\{W\}) \Psi_n(\{w\}; \{W\}) \quad \hat{H}_n \chi_n(\{W\}, t) = i\hbar \frac{\partial}{\partial t} \chi_n(\{W\}, t)$$

+ Statistical averaging for beam conditions



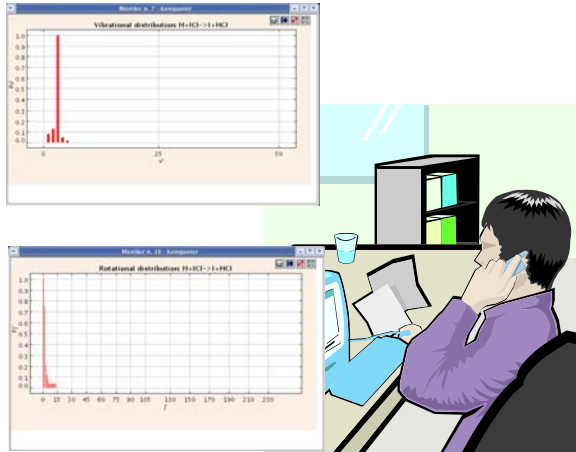




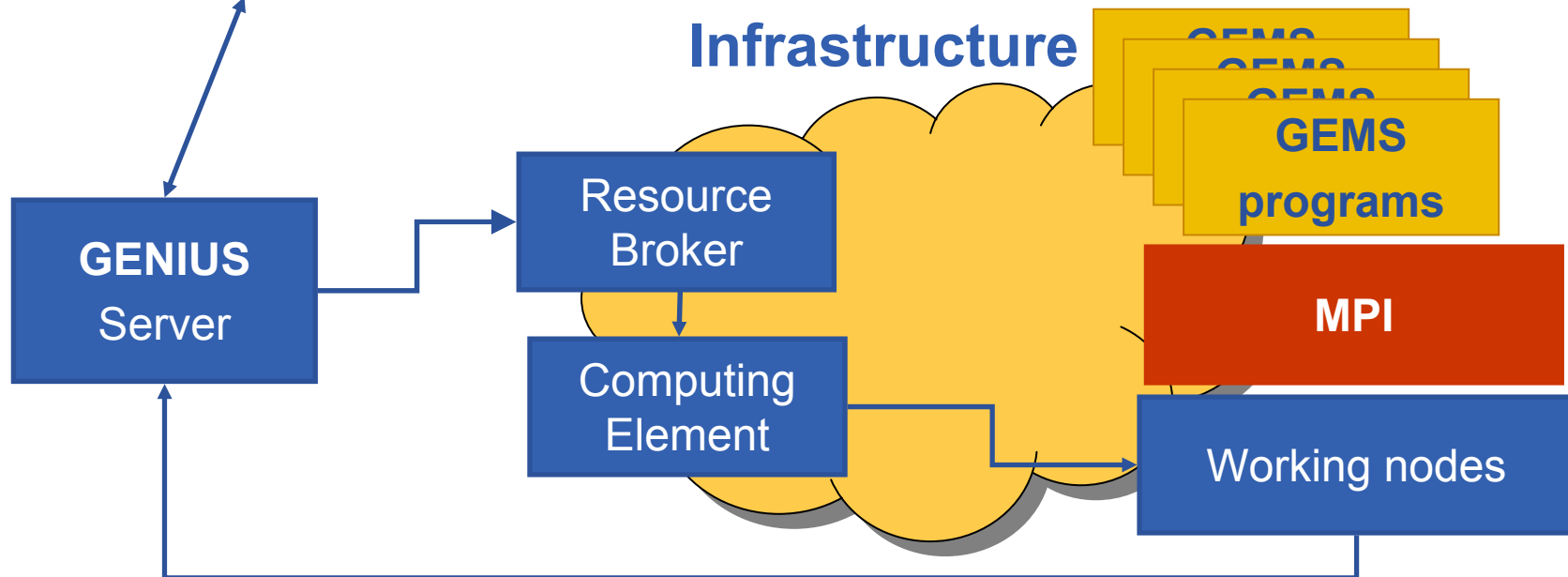


- **Sites**
 - GILDA testbed sites, in which INFN Grid 2.2.0 (fully compatible with LCG 2.2.0) middleware is installed.
- **Key services**
 - Resource Broker: `grid004.ct.infn.it`
 - Computing Element: `ce.grid.unipg.it`
 - User Interface: `ui.grid.unipg.it`
 - GENIUS Portal: <https://genius.ct.infn.it>
 - CompChem VOMS

- **Number of jobs: 10 per day**
- **Storage: 100KB – 50GB depending on the type of computational engine used and the chemical system studied:**
 - Trajectory calculations: <100KB
 - Time Dependent Quantum: 10GB
 - Time Independent Quantum: 50GB
- **RAM: 100KB – 2GB**
 - Trajectory calculations: <100K
 - Time Dependent Quantum: 1GB
 - Time Independent Quantum: 2GB
- **Success rate: 98%**



GILDA Testbed Infrastructure



Outbound connectivity

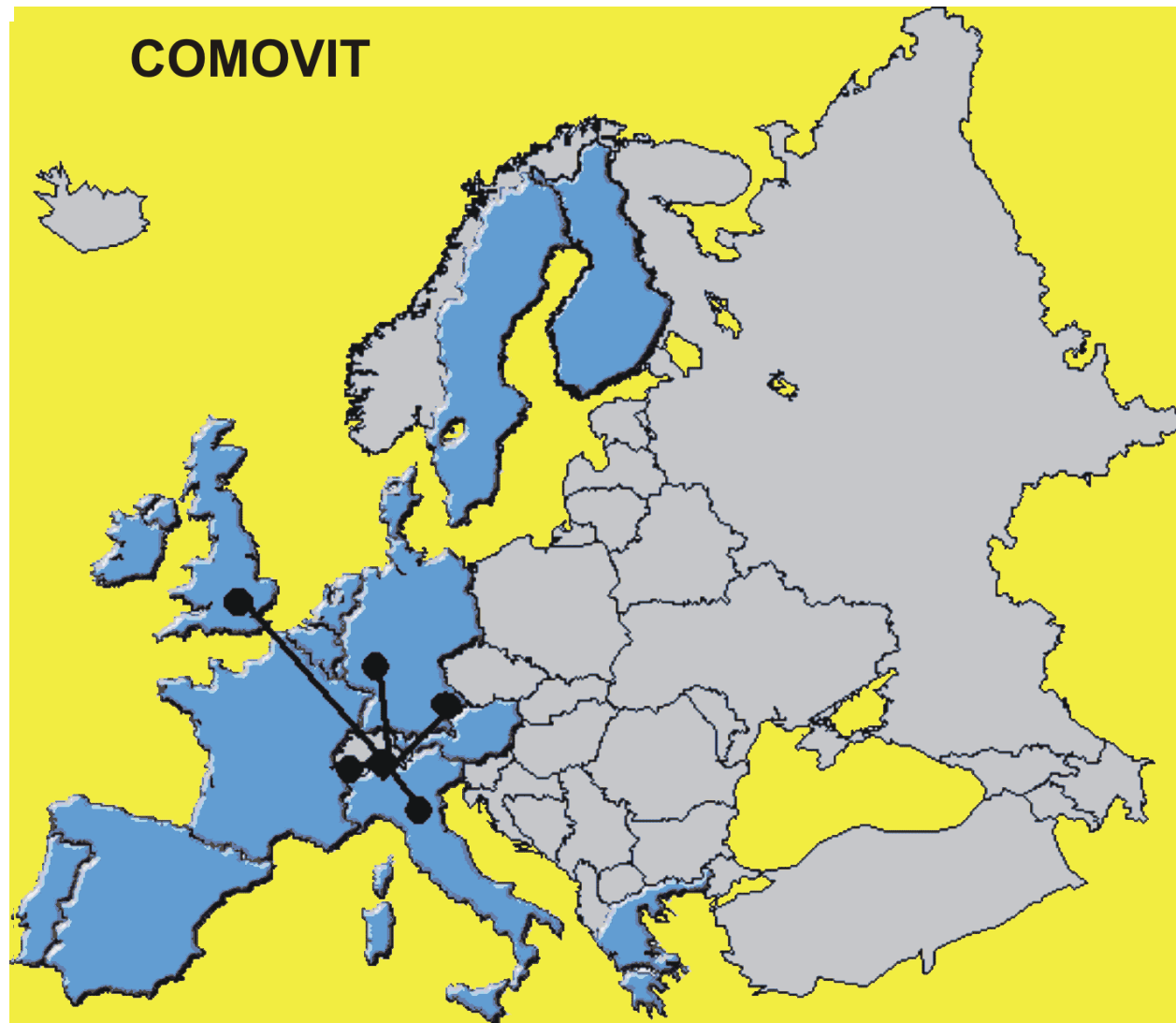
- **We have encountered the following problems:**
 - Not all sites did publish the environment variables related to the outbound connectivity of the Working Nodes.
 - The job activation time is high, especially if an high number of working nodes is requested.
 - Not all sites have enough resources (disk + RAM) for Quantum Chemistry codes.
 - In some sites MPI is not available.
 - The Firewall sometimes is misconfigured and blocks the outbound connectivity in the EGEE standard port range: 20000-25000.
- **The workarounds were**
 - Alert people involved to properly configure the environment variables.
 - Alert people to open the EGEE standard ports on the Firewall.

- **Running at weekends computers of teaching laboratories (trajectories).**
- **Making a combined use of SP4 at CINECA and CRAY T3E at Edinburgh (3D quantum).**

- **What's missing to extend the application to production**
 - Define the policy for the usage of commercial suites of programs. At present each site buys its own license and opens the access to the VO users. The commercial licence requirement for a job need to be specified in a JDL statement
 - Credits for contributions to development and service
 - Extend the suites of programs to other components of the molecular simulator
 - Full deployment of the VO and consortium
- **As Production Service, the expected usage profile is:**
 - Users: 100
 - Sites: 30
 - CPU consumption (per job): 100 - 1000 hours
 - Disk Storage (per job): 10 MB - 800 GB
 - RAM Storage (per job): 100 KB – 200 GB



- Simbex
- Murqm
- Dirac
- Elchem
- Icab
- Dysts
- Comovit



- 1 Isr,Pl,Sk,NI
- 2 Cz,Ch, Fr, Dk, A, Sw, No
- 3 Hu
- 4 Gr
- 5 E
- 6 D, Uk,
- 9 I

- **3 Computer Centres**
- **New electronic structure programs (MOLCAS, DIRAC, DALTON, COLUMBUS, MR-CCSD).**
- **New Dynamics programs (AMD, TPS, KMC, condensed phase).**
- **Chemical knowledge semantic web (molecular structures, apparatuses, processes).**

- **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)
- **Computational Campaigns**
 - to evaluate properties depending on the fate of few out of millions, billions or more events by distributing the execution of the computations on the grid (in the prototype particular reactive events having specific energetic or steric properties)
- **Collaborative Engineering of knowledge**
 - to handle (when is the case also in a privacy protecting fashion) chemical information and knowledge including training and production of new knowledge (for the prototype the design of virtual chemical processes or laboratory experiments)

- **A consortium agreement will be signed by the new participants of CompChem VO.**
- **A dedicated Portal for GEMS will be implemented, based on MyProxy authentication.**
- **Various computational engines will be implemented in order to offer grid services based on the a priori simulation of molecular systems.**
- **Molecular Virtual Reality components will be combined with Human-size Virtual Reality to offer multiscale realistic representations of material structures and processes.**
- **Design appropriate cost models for supporting research, project and service activities.**