Use Case of gLite Services Utilization

Multiple Ligand Trajectory Docking Study

Jan Kmuníček

CESNET NA4 & JRA1 effort

www.eu-egee.org

Information Society

- **Scientific context**

- **Grid added value**

- **Implementation details**

- **Achievements & plans**

- **Introduction to the application**

  - interactions between large biomolecules and smaller bio-active ligands

    - lies on the **foundation of many biological properties**
    - interest in the biomolecular/pharmaceutical research

  - interactions studied through a process called **docking**

    - aims to find energetically favorable orientation of the **ligand** within an active site of a biomolecule

eGee

Enabling Grids for E-sciencE
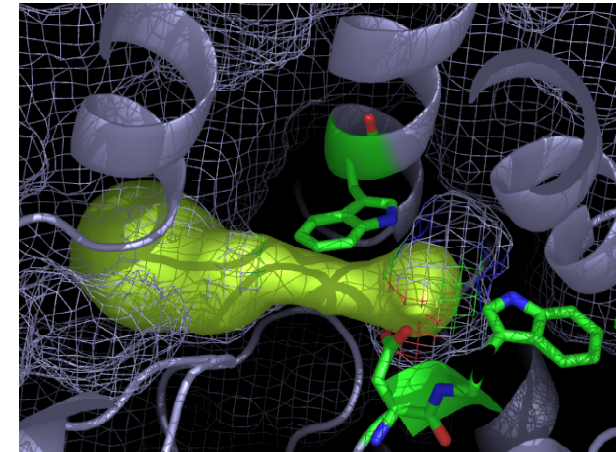
- **Introduction to the application**

  – search is done on snapshots taken
    from molecular dynamic trajectory
    describing dynamic behavior
    of a biomolecule and a set of
    candidate ligands
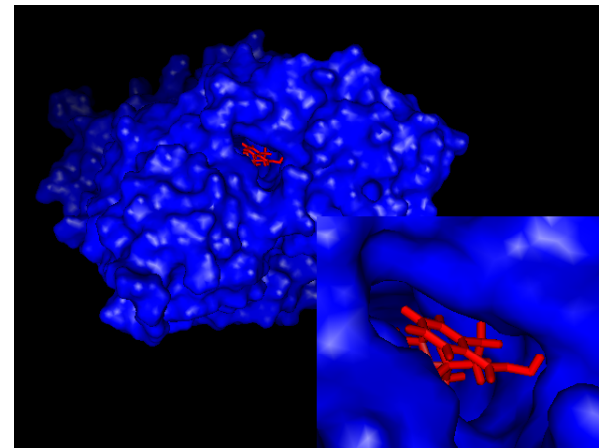
  – such computation is **infeasible without Grid infrastructure**
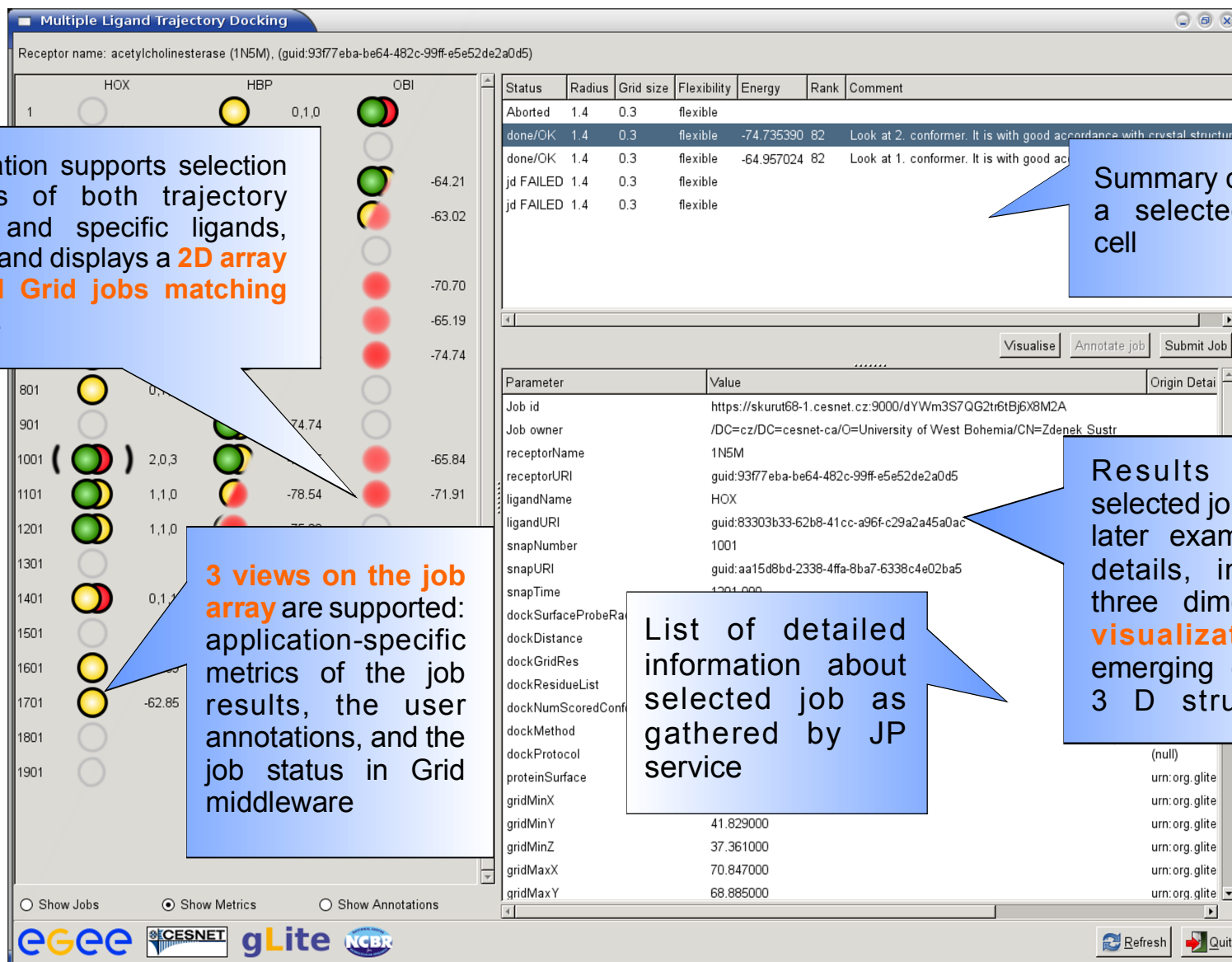
  – Grid environment serves for

    performing required computations   **job submission**
    management of results   **provenance service**

- **Biomolecular aspects of chemical weapons**

  – many **organophosphate nerve agents**

  – interference with the signaling system used by the nervous system

  – these compounds inhibit acetylcholinesterase (AChE) essential chemical that **breaks down nerve signals between nerve cells**

  – looking for universal reactivator able to liberate AChE when it is poisoned by the nerve paralytic compounds

eGee

Enabling Grids for E-sciencE

- **Added Value for Application Experts**

  ✓ GUI application for solving generic biomolecular parametric jobs

  ✓ user application metrics evaluation based on targeted parameters

  ✓ potential extension for biomedical screening

  ✓ effective collaboration among researchers

  ✓ computational jobs manipulation (input modification, jobs resubmission)

  ✓ targeted search & selection of desired jobs (finished / non-finished, aborted)

  ✓ modular application specific plugins for results presentation (*e.g.* visualization)

The application supports selection of subsets of both trajectory snapshots and specific ligands, queries JP, and displays a **2D array of finished Grid jobs matching the criteria**.

Summary of jobs in a selected array cell

3 views on the job array are supported: application-specific metrics of the job results, the user annotations, and the job status in Grid middleware

List of detailed information about selected job as gathered by JP service

Results of the selected job can be later examined in details, including three dimensional **visualization** of emerging complex 3 D structures.

Enabling Grids for E-sciencE
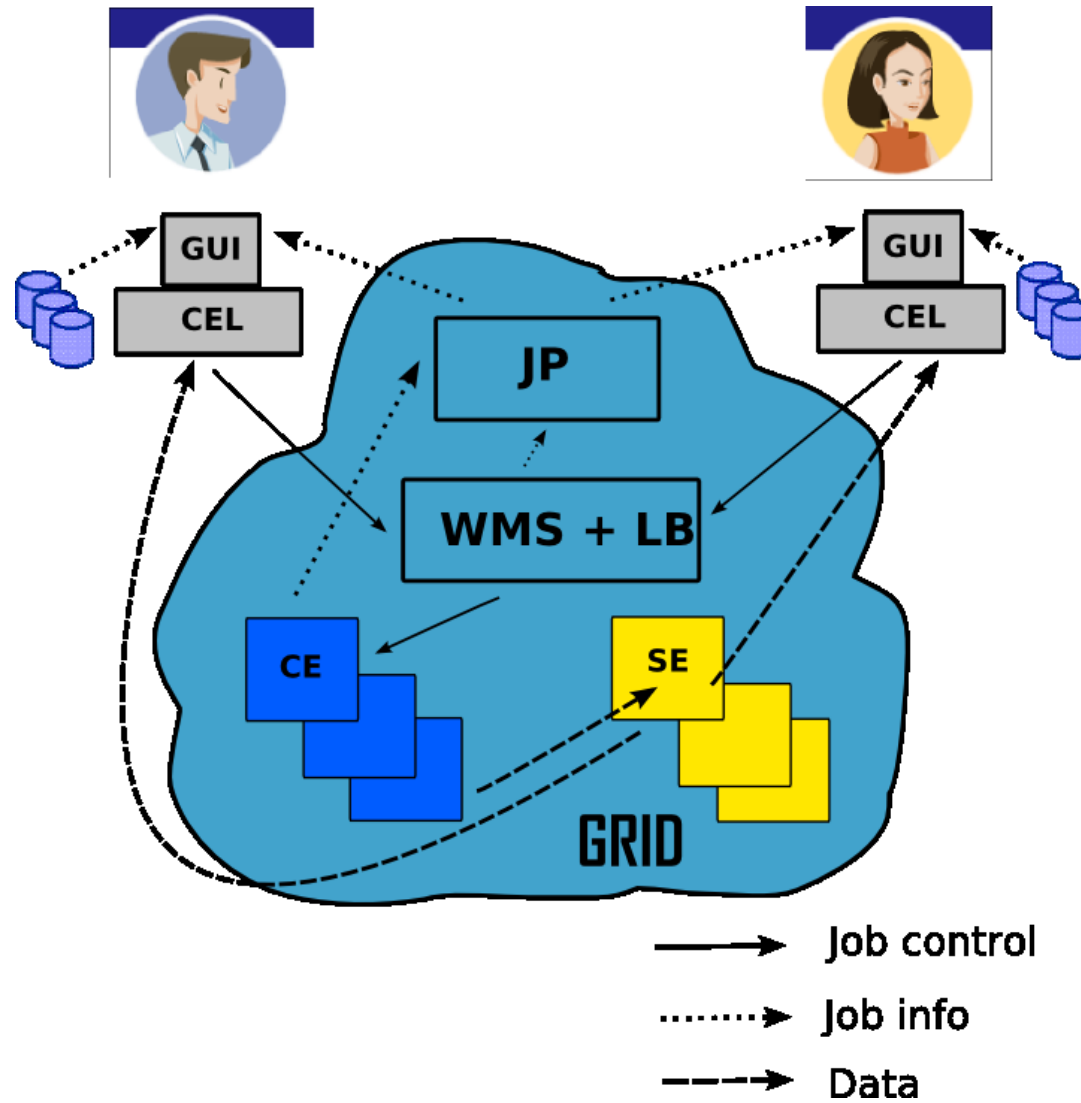
- ## Job Provenance

  - **Job Provenance Motivation**

    - the information about jobs has longer value
      - *e.g. repeat a submission of a job executed year ago*

    - the information about job control flow and job execution environment complements job results
      - *e.g. to be able to reliably resubmit a job*

  - **Job Provenance Service**

    - preserve information about Grid jobs
    - allow data-mining in this information
    - assist jobs re-submission

- **Charon Extension Layer**

  – **Charon Extension Layer (CEL) Motivation**

    ▪ uniform and modular approach for (complex) computational jobs submission and management

    ▪ generic system for use of application programs in the Grid environment (LCG/gLite middleware, …)

  – **Charon Extension Layer Service**

    ▪ application management
      - *single/parallel execution without job script modification*

    ▪ job management
      - *easy job submission, monitoring, and results retrieval*

- **Virtual Organization for Central Europe (VOCE)**

    – provides **complete grid infrastructure** under EGEE wings

        ▪ VOCE officially registered in the list of EGEE VO
        ▪ the first regional VO in EGEE
        ▪ currently this model is recommended for all regions

    – based on **regional principle**

        ▪ VOCE spans the whole Central Europe (CE) Federation
        ▪ core services operated by CESNET
        ▪ resources are provided by many institutions across the CE
          (these resources are available to all users registered in VOCE)

    – documentation   http://egee.cesnet.cz/en/voce/index.html

Enabling Grids for E-sciencE

- **Virtual Organization for Central Europe (VOCE)**

  resources from   Austria

                                   Croatia

                                   Czech Republic

                                   Hungary

                                   Poland

                                   Slovakia

  more than   **283 registered users**

                                   from 14 institutes and 6 countries

  in total   **1248 CPUs**   **(18 CEs)**

                                 cca 32.6 TB disk space   **(17 SEs)**

- **Advantages**

  – manual work with results (annotation, comparison, evaluation) easily substituted by semi-automatic tool supporting **in-depth overview of finished jobs**

  – support typical workflow for day-to-day work

  – effective collaboration among researchers highly improved

  – simplified **integration of research newcomers** into studied scientific problem

  – leveraging Grid services while hiding Grid complexity

- **Suggested enhancements**

  – allow not only Grid jobs output annotations but job data item annotations too

  – extend the functionality towards **generic molecular dynamic trajectory selection** (this feature currently under development)

- **Uncover all details!**

  – visit available Wiki pages at

    http://egee.cesnet.cz/mediawiki/index.php/Job_Provenance_Demo