Molecular Modelling of Complex Systems with InSilicoLab

Joanna Kocot, Andrzej Eilmes,
T. Szepieniec, M. Sterzel, and M. Golik

ACC CYFRONET AGH,
Faculty of Chemistry, Jagiellonian University.

InSilicoLab: Idea

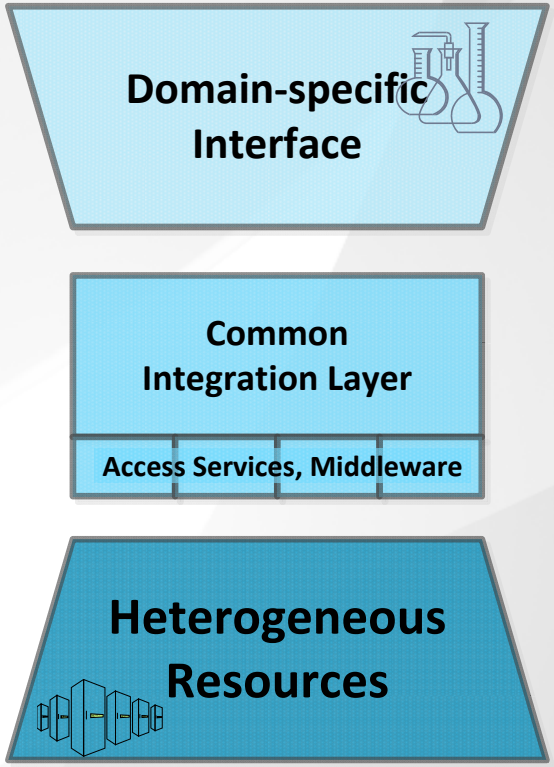
- Workspace gathering all that a researcher needs for their *in silico* experiments
  - Enabling performing large-scale, long-lasting data- and computation-intensive experiments
  - Facilitating categorisation and description of data
    - Enabling searches and browsing

Method: Utility

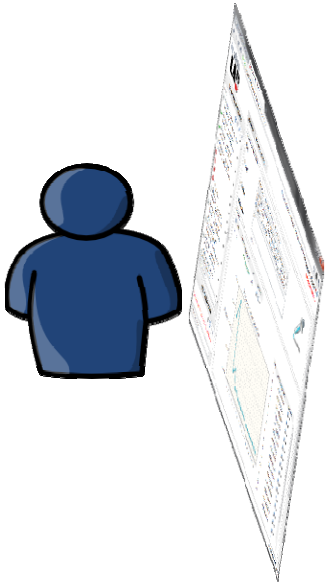
- Analysing the researchers work:
  - Ways of working
  - Common problems
- We intend to aid in solving **SPECIFIC** problems
- We cannot support solving **ANY** scientific problem
  - We assume it cannot be done in a universal and comprehensive way

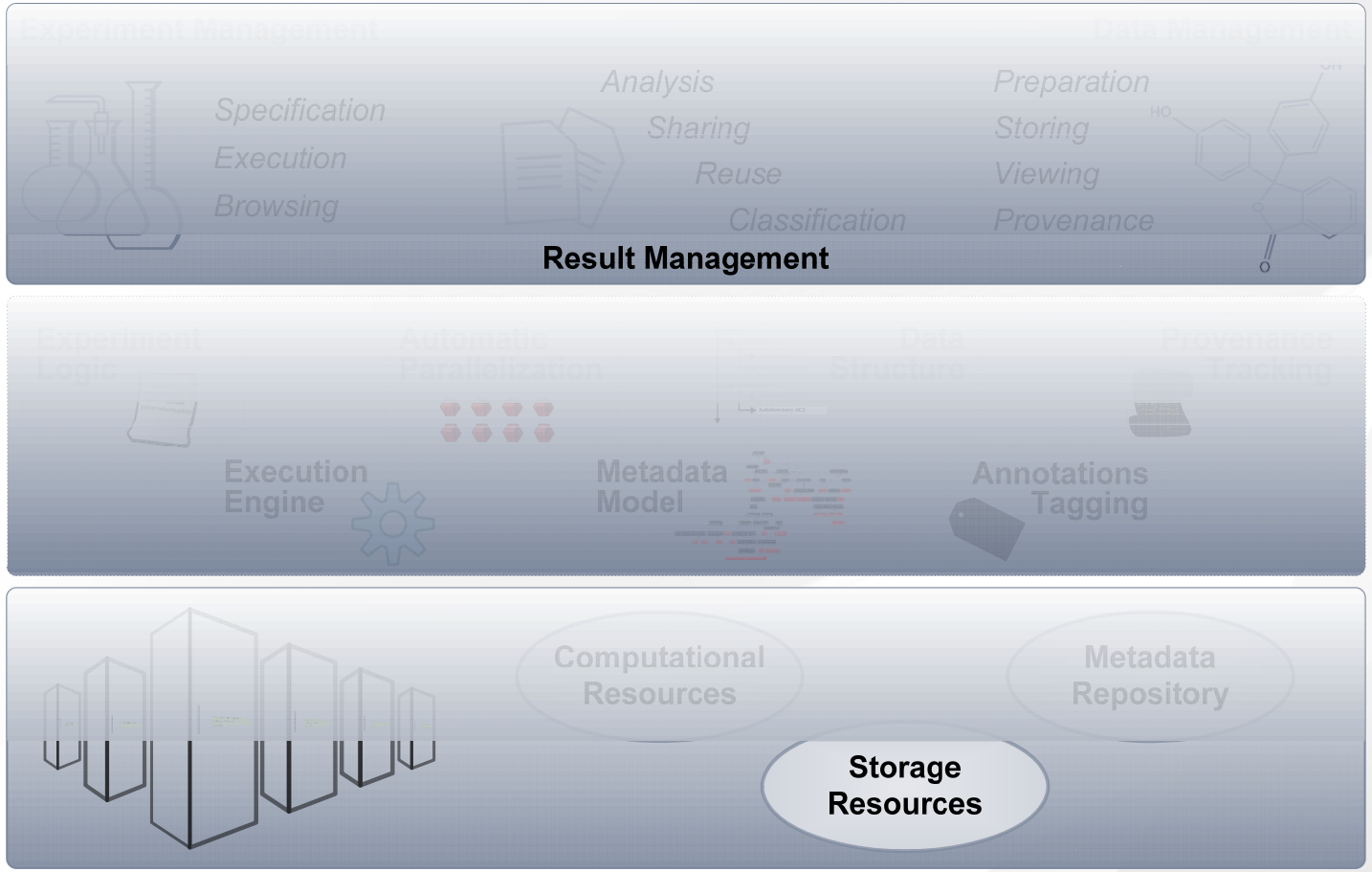
Architecture

- System should be optimised for problems with common characteristics, like:
  - Large resource consumption
  - Repeatability – experiments conducted in similar way
- It is not limited to these problems
- Architecture of the whole system is generic
  - Ensures access to large, heterogeneous computing and storage resources
  - Through integration layer – built on top of resource access services, middleware, etc.
  - Presented to the user with domain/problem-specific interface



InSilicoLab Architecture – Details





Domain Layer
Managed by the user

Mediation Layer
translation from domain- to resource-specific language

Resource Access Layer
access to e-infrastructures

How we create/integrate new experiments

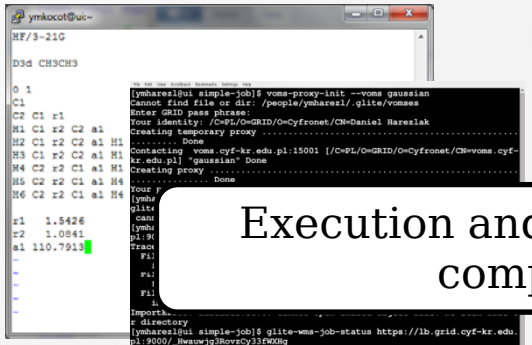
- Discover a pattern in the researchers work
  - A joint effort of the developers and the researchers teams
- Put it down as an algorithm – experiment logic
- Translate into necessary scripts
  - Include input and results management
  - Allow metadata attachment
- Adjust interface
  - Input specification
  - Result display
  - If neccessary: new data types management

QC/MD Modelling

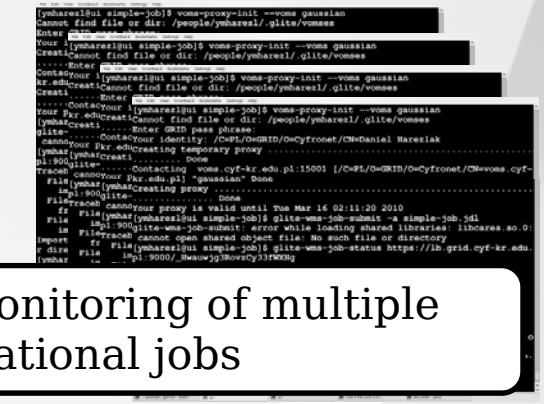
- Quantum-chemical modeling of complex systems often requires:
  - Analysis of large molecular systems
  - Preparation of large sets of input files sharing common pattern
  - Execution of multiple computational jobs
  - Postprocessing of output files

QC/MD Modelling

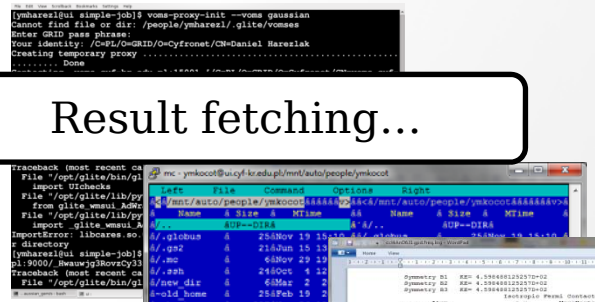
Preparation of input files

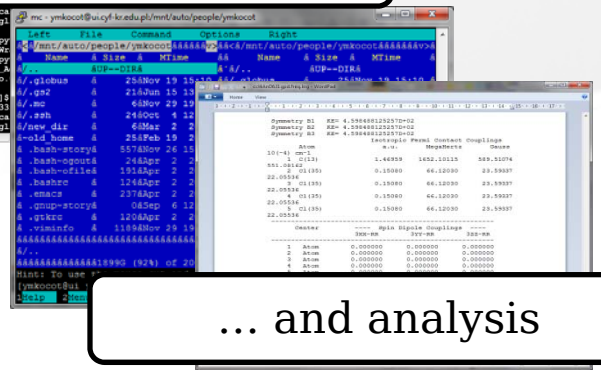


Execution and monitoring of multiple computational jobs

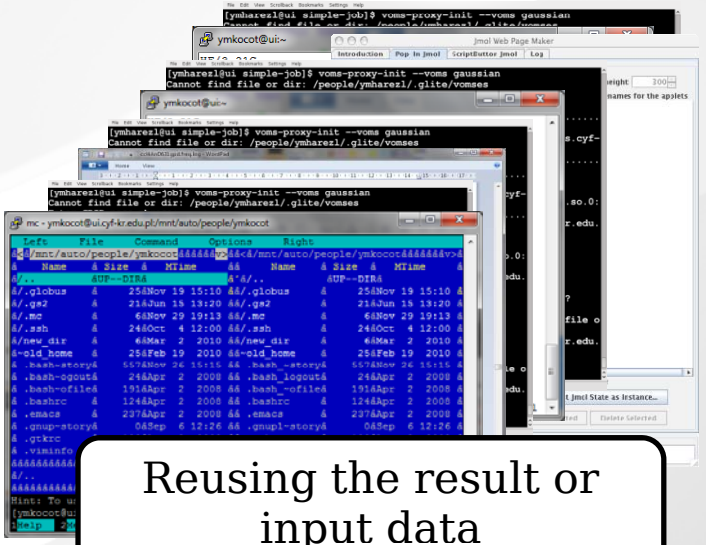


Result fetching…





… and analysis

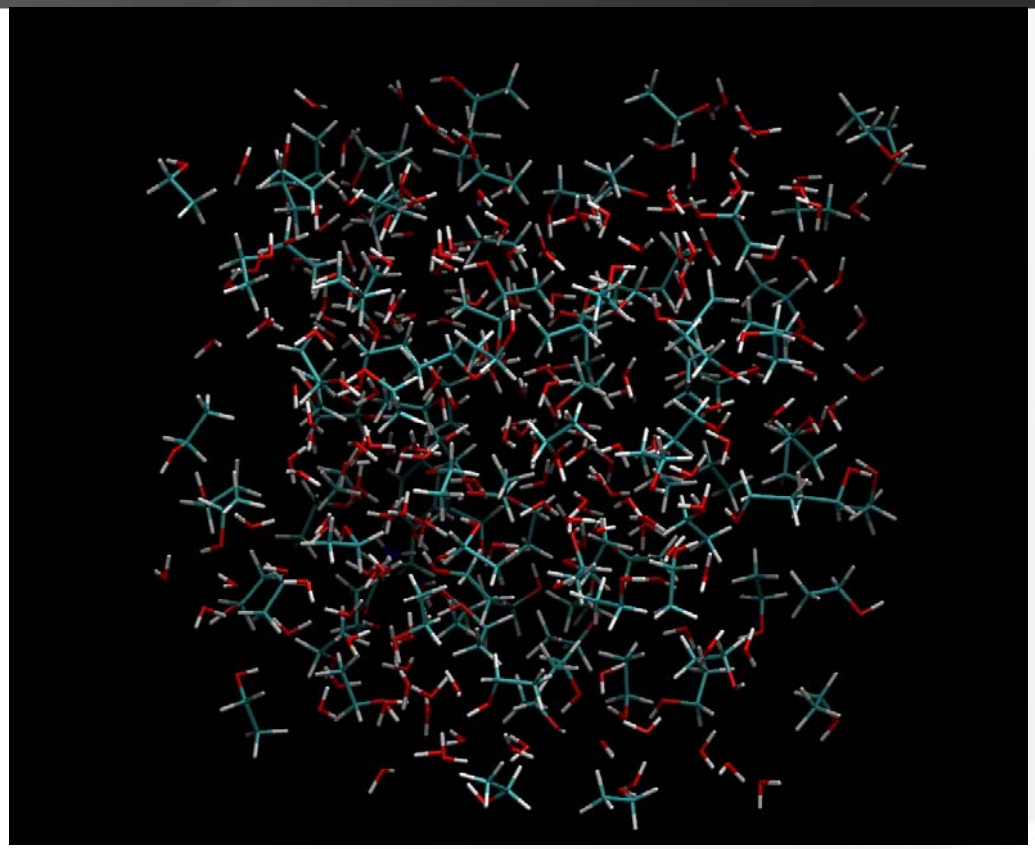


Reusing the result or input data

QC/MD Modelling: Sequential Approach

- Sequential approach to molecules in solution:
  - Simulation of the solute-solvent system at the Molecular Dynamics level (classical or ab initio)
  - Selection of the solute with its solvation shell and performing quantum-chemical calculations (possibly at higher level of theory) to calculate solvent effect on physical properties (energies, excitation energies, chemical shifts)

What is more... Reduction



Reduction of the system necessary to:

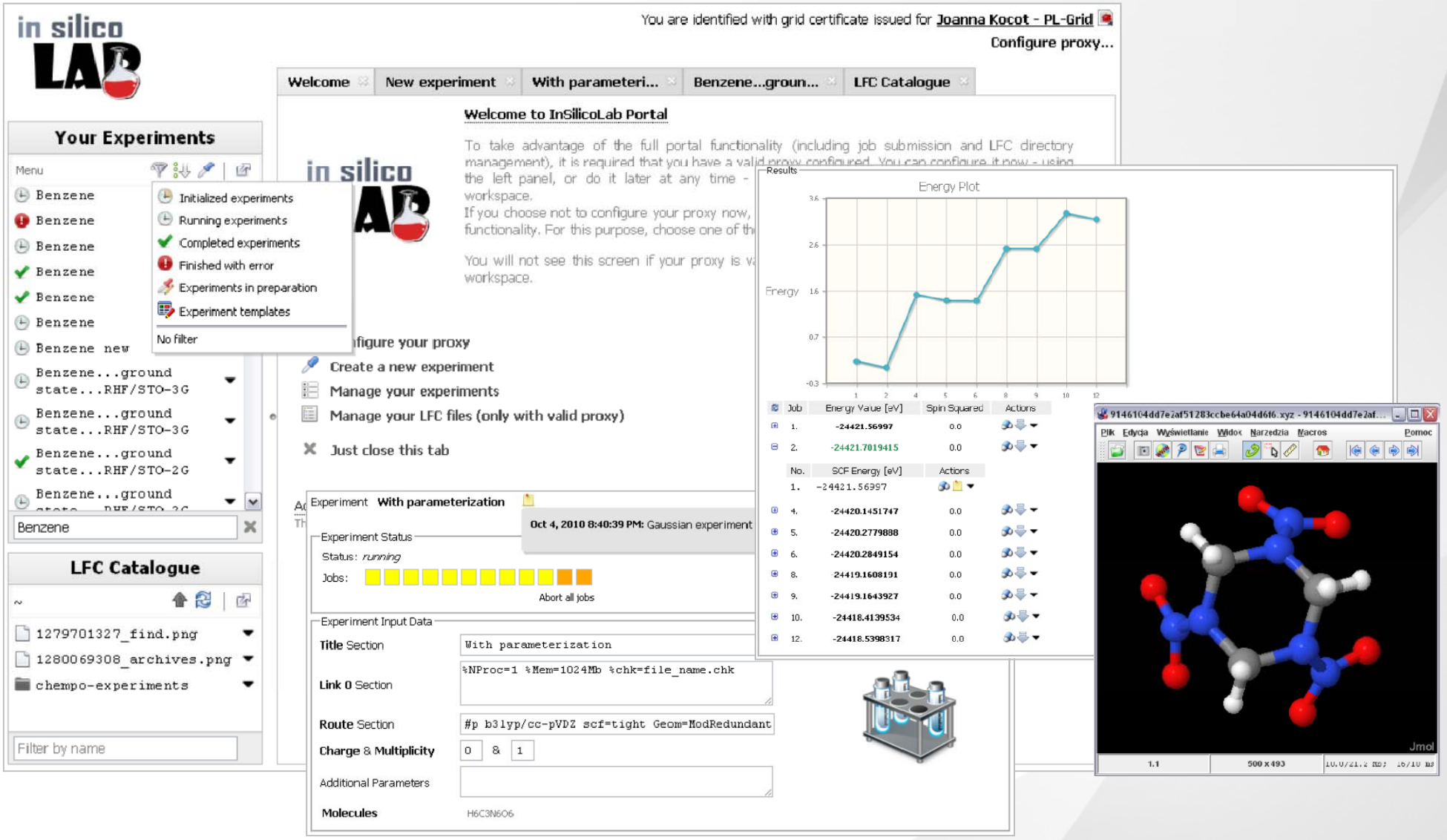
- Visualize the „reaction center”
- Make the quantum-chemical calculations feasible

Requires from the researcher additional effort and usually some programing skills

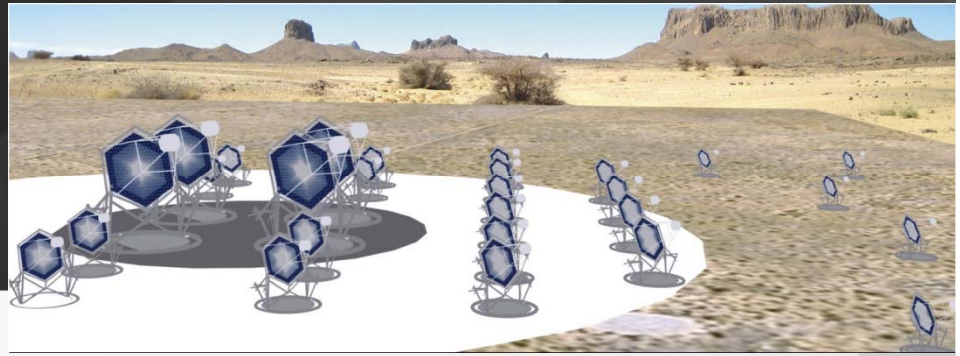
QC/MD Modelling: Desired Solution

- Desired features of a tool facilitating moecular modelling data manipulation:
  - Applicability to a variety of systems
  - Easy and flexible definition of the selectable part of the system
  - Visualization at each step
  - Automated preparation of series of geometries and input files
  - Execution of QC jobs on the available infrastructure (PLGrid)
  - Results fetching
  - Parsing of the results and performing basic statistics

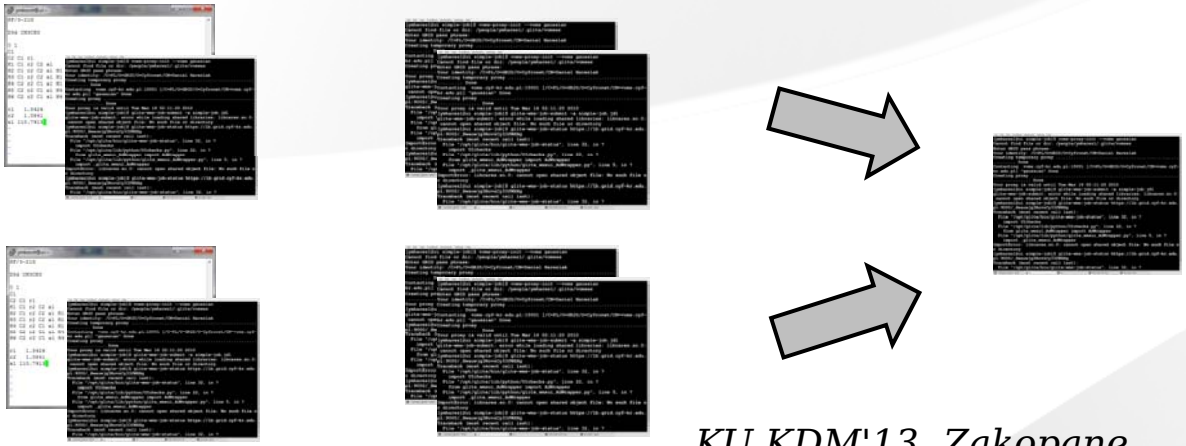
InSilicoLab for Chemistry

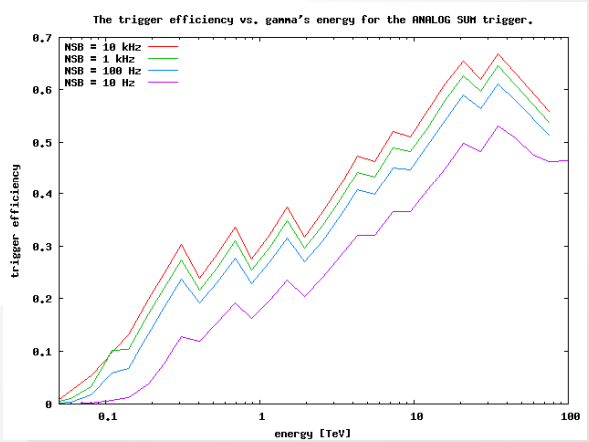


Other domains



- Cherenkov Telescope Array (CTA)
  - Project for ground based gamma-ray astronomy
  - Large international collaboration (25 countries)
  - Scientific goal is to build telescope array:
    - 10 times more sensitive than current instruments
    - with better energy and angular resolution
    - with wider field of view and energy coverage
    - with budget ~ €190 M





Other domains

- MHD code created at Centre for Astronomy, Nicolaus Copernicus University in Toruń, Poland
- Integration with InSilicoLab done by the Piernik developers
  - Only aided by the InSilicoLab team
- Will be deployed as PL-Grid service
- http://piernik.astri.umk.pl/

Summary

- The *InSilicoLab* portal is a solution for researchers performing *in silico* experiments in many domains of science
  - Specific problem support, but supports many problem classes – thanks to generic system architecture
- Validated and running in production mode for two domains (one beta-stage deployment)
  - Will be released as PL-Grid service in March (Chemistry and in December (Piernik)
- Open for new collaborations

Thank you!

http://insilicolab.grid.cyfronet.pl

insilicolab@cyfronet.pl