Comprehensive Support for Chemistry Computations within PL-Grid Plus and PL-Grid NG Projects

A. Eilmes, J. Kocot, M. Makowski, G. Mazur, K. Noga, M. Sterzel
Computational chemistry within PL-Grid

Outline

1. Introduction
2. InSilicoLab Framework
3. Services
4. Summary
Computational chemistry nowadays

Computational and quantum chemistry

- made it possible to predict and explain experimental results
- became valuable methodology in material science, biochemistry and related areas

but modern modeling

- incurs huge computational effort
- requires proper choice of adequate and efficient methods
- produces large sets of data which need to be stored
Computational chemistry nowadays

Computational and quantum chemistry
- made it possible to predict and explain experimental results
- became valuable methodology in material science, biochemistry and related areas

but modern modeling
- incurs huge computational effort
- requires proper choice of adequate and efficient methods
- produces large sets of data which need to be stored
Call for integrated solutions

Existing ecosystem of computational tools
- is fragmented
- lacks compatibility of its parts

We build a platform for computational chemists that
- promotes efficient use of existing infrastructure
- support researchers in using most adequate tools
Computational chemistry within PL-Grid

InSilicoLab Framework

Overview

http://insilicolab.cyfronet.pl

- Framework of application portals
  - enables execution of large-scale, long-lasting, data-intensive numerical experiments
  - supports management of complex calculations
  - enables joint analysis of results from massively parallel computations

- Scientific domains
  - quantum chemistry
  - Cherenkov Telescope Array (CTA)
  - multihydrodynamic for astrophysics
Computational chemistry within PL-Grid

InSilicoLab Framework

Technology

- Server
  - Java (jSAGA)
  - Ruby
- Computing nodes
  - Domain-specific software
  - Python
- Middleware
  - gLite
  - DIRAC
- Web portal
  - Google Web Toolkit
  - Jmol
  - JSMol

Preliminary version – 28 października 2014
Chemistry

- PL-Grid Plus
  - General Quantum Chemistry Experiment
  - Trajectory Sculptor
  - Cubegen
  - QCAdvisor
- PI-Grid NG
  - GPGPU Support
  - MOOSE Database
Workspace provides assistance with

- preparation of input files for various QC applications
- performing calculations on grid infrastructure
- monitoring and controlling complex numerical experiment workflow
- collecting output data
- categorization and visualization of data
- analysis of the numerical experiment result
Computational chemistry within PL-Grid

Services

General Quantum Chemistry Experiment

More detailed look

Integrated with:

- Gaussian
- Turbomole
- GAMESS
- Niedoida
Molecular modeling of complex systems often requires a two-stage approach

- Molecular Dynamics (MD) simulations
  - to generate possible structures of the system
  - to investigate evolution of the system in time

- QC calculations
  - for the selected parts of the system
  - to calculate desired properties at higher level of theory
  - possibly for the series of frames from the MD trajectory

Trajectory Sculptor is a tool to extract relevant parts of a MD trajectory, which can be used subsequently in General Quantum Chemistry Experiment environment.
Three-stage trajectory reduction
- system definition: user specifies composition of the system and defines solute and solvent
- distance definition: distance metric is supplied and distance threshold for selection of the relevant part of the system is set
- choice of frames

Preview of the results at each stage

Support of various data formats

Support for periodic and open boundary conditions

Flexibility of selection criterions
Computational chemistry within PL-Grid

Services

Trajectory Sculptor

Sample sculpture
Computational chemistry within PL-Grid

Services

GPGPU Support

GPGPU

GPGPU in chemistry

- High efficiency
- Low energy consumption
- Accessible

but

- Specific software slows adoption
GPGPU in chemistry

- High efficiency
- Low energy consumption
- Accessible

but

- Specific software slows adoption
InSilicoLab for Chemistry – GPGPU support

- InSilicoLab as common platform
  - Workflow (mostly) independent from specifics of particular program
  - Lower entry barrier for new computational packages

- Issues with GPGPU support
  - Restricted middleware
    - Already (at least partially) solved
  - Required support for new quantum-chemical packages
    - TeraChem
    - adorycja
    - ADF (perhaps)

is in progress
Computational chemistry within PL-Grid

Services

GPGPU Support

InSilicoLab for Chemistry – GPGPU support

- InSilicoLab as common platform
  - Workflow (mostly) independent from specifics of particular program
  - Lower entry barrier for new computational packages

- Issues with GPGPU support
  - Restricted middleware
    - Already (at least partially) solved
  - Required support for new quantum-chemical packages
    - TeraChem
    - adorycja
    - ADF (perhaps)

is in progress

Preliminary version – 28 października 2014
Intended to provide users with information about the implementation of various QC methods in the software packages available in PL-Grid infrastructure. It consists of

- underlying knowledge base summarizing the capabilities of QC software available
- user interface in form of wizard guiding an user
  - from the definition of the problem (molecular system and properties wished)
  - to the input files prepared for a specific software package chosen as most promising to provide satisfactory solution in acceptable time.
Intended to provide users with information about the implementation of various QC methods in the software packages available in PL-Grid infrastructure. It consists of:

- underlying knowledge base summarizing the capabilities of QC software available
- user interface in form of wizard guiding an user
  - from the definition of the problem (molecular system and properties wished)
  - to the input files prepared for a specific software package chosen as most promising to provide satisfactory solution in acceptable time
Computational chemistry within PL-Grid Services

QCAdvisor

(Mini) expert system

Intended to provide users with information about the implementation of various QC methods in the software packages available in PL-Grid infrastructure. It consists of

- underlying knowledge base summarizing the capabilities of QC software available
- user interface in form of wizard guiding an user
  - from the definition of the problem (molecular system and properties wished)
  - to the input files prepared for a specific software package chosen as most promising to provide satisfactory solution in acceptable time
Intended to provide users with information about the implementation of various QC methods in the software packages available in PL-Grid infrastructure. It consists of

- underlying knowledge base summarizing the capabilities of QC software available
- user interface in form of wizard guiding an user
  - from the definition of the problem (molecular system and properties wished)
  - to the input files prepared for a specific software package chosen as most promising to provide satisfactory solution in acceptable time
Intended to provide users with information about the implementation of various QC methods in the software packages available in PL-Grid infrastructure. It consists of

- underlying knowledge base summarizing the capabilities of QC software available
- user interface in form of wizard guiding an user
  - from the definition of the problem (molecular system and properties wished)
  - to the input files prepared for a specific software package chosen as most promising to provide satisfactory solution in acceptable time
Computational chemistry within PL-Grid

Services

QCAdvisor

Prototype

Diagram showing the process flow:
1. Start
2. Definition of molecular system
3. Desired properties
4. Suggested computational method
5. Proposed basis set
6. User defined basis set
7. Proposed computational package
8. Input file(s) for selected computational package
Computational chemistry within PL-Grid

Services

MOOSE Database

Purpose

- data source for benchmarking existing and newly proposed QC methods
- assistance in making an educated choice of QC method appropriate for modelling a given phenomena
- data source for expert systems in InSilicoLab for Chemistry
Computational chemistry within PL-Grid

MOOSE Database

Data

Structural and spectroscopic data from QC calculations and directly comparable high-quality experiments results

- adiabatic ionization potential and electron affinities
- adiabatic electronic excitation energies
- geometrical structures of ground and excited states
- vibration frequencies in ground and excited states

Sources

- literature (experiment/calculations)
- local (calculations)
- social: future users (possibly including InSilicoLab experiment results)
Features

Service should allow for

- searching the data according to user-defined criteria
- aggregation and statistical analysis of the search results
- adding new data and making critical evaluation of the existing one
Take-home messages

- several projects are run with the aim to create user-oriented environment for computational chemists accessing PL-Grid infrastructure
- a few new services (QCAdvisor, GPGPU support, MOOSE database) are currently at the (pre-)prototype stage
- with more than 100 users registered for InSilicoLab for Chemistry usefulness of developed solutions already proved