

Current Status of Chempo Web Portal

M. Sterzel, T. Szepieniec, D. Harężlak
ACC Cyfronet AGH

EGEE'09,
21-25 September 2009, Barcelona, Spain

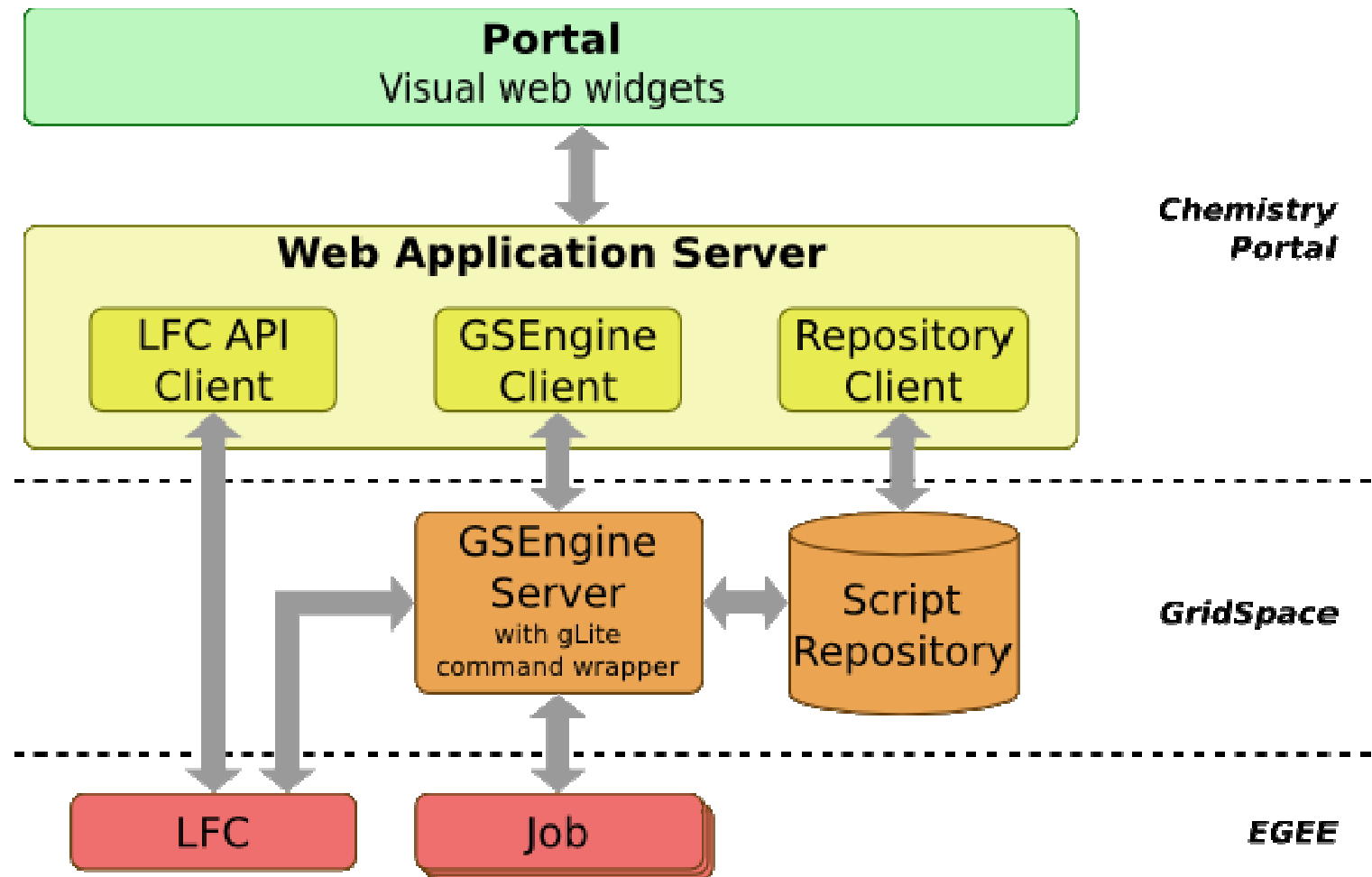
- **Grid application in chemistry**
- **Aim of the project**
- **Technology**
- **API details**
- **Metadata description**
- **Sample screenshots**
- **Future plans**

- **Tasks to which Grid can be applied directly:**
 - Conformational analysis
 - Numerical frequency computations
 - Zero Point Vibrational Averaging
 - Property computations for series of geometries from Molecular Dynamics simulation
 - Determination of chemical reaction paths
 - Determination of potential energy surfaces (PES)
- ... all kind of “brute force” tasks, or tasks which operate on huge data sets
- **A common factor – set of molecular geometries for which we repeat the same computations**

... to build an environment for chemists with integrated tools which:

- enable convenient and natural management of computational processes and experiment data
 - without changing chemist's way of thinking
 - and not being distracted by technology
 - enables cooperation between scientists
-
- **Comparing to any other grid portal our product is experiment-centric not grid-centric**

- **Web 2.0 project and all other means to make portal convenient for users**
- **Build with help of Google Web Toolkit**
- **API for application, mapping, submission, data validation, processing, management, etc.**
- **GForge to enable team collaboration**
- **Metadata to help manage files in grid storage**
- **Extensive use of tools developed by ViroLab project**
 - Experiment Planning Environment (EPE)
 - GSEngine for job management



- **The LFC API Client is used to manage the user data being results of the application executions. Operations available:**
 - retrieve user data items
 - retrieve the contents of the file
 - retrieve partial file content
 - store file's metadata
 - retrieve file's metadata
- **GSEngine Client for job managing**
- **Repository Client to store experiments**
- **VOMS proxy API – used to communicate with the VOMS server in order to generate user proxies**

- **Single directory for each experiment**
 - file `.metadata` includes
 - description
 - who run experiment
 - when it was run
 - directory `input`
 - no metadata
 - file `experiment_name` that includes all parameters to rerun experiment
 - directory `output`
 - file `.metadata` includes all metadata concerning files in directory
 - *each metadata:*
 - file (string)
 - parameter (string)
 - value (double)
 - description (string)

- Written in Ruby language
- Describe the whole experiment from data preparation through job execution and output data postprocessing

```
require 'cyfronet/gridspace/goi/core/g_obj'
require 'cyfronet/gridspace/goi/utils/lcg/job_spec'
require 'rexml/document'

def gaussianJob(gcount, gExec, gGeom, gHead, gTail)

  printf("Executing experiment for %d geometry\n", gcount)
  arguments=gExec + " -c " + gcount.to_s + " -g " + gGeom
  arguments+=" -h " + gHead + " -f " + gTail

  # Create JDL specification here
  jobSpec = JobSpec.new
  jobSpec.executable='/bin/bash'
  jobSpec.arguments=(arguments)
  jobSpec.stdoutoutput='stdout.out'
  jobSpec.stderr='stderr.err'
  jobSpec.add_to_output_sandbox('stdout.out')
  jobSpec.add_to_input_sandbox(gGeom)
  ...
end
```



Experiment - DynamicKaplanMeier/src/DynamicKaplanMeier.rb - ViroLab EPE

File Edit Source Refactor Navigate Search Project Window Help

Experiment Resour

DynamicKaplanMeier 735 [Repo: src 682]

DynamicKaplanMeier.rb 6

DataKaplanMeier.csv 588

.project 735

experiment.xml 406

ExperimentLicense.txt 406

feedback.xml 548

log.txt 406

log.txt.1 406

```
## DynamicKaplanMeier.rb
require "cyfronet/gridspace/goi/core/g_obj"
require "cyfronet/gridspace/dac/DACConnectC
require "cyfronet/gridspace/ulib/FormDataRe

# request file upload from a user
input = GS_CLIENT_CALLBACK.getFileAsString("
#puts input

# send file to a dav server
dav = DACConnector.new("dav","https://viro
dav.storeResult(input, 'kap-meier')

# get column names from csv file
# first line of input
header = input[./.$$/]
columns = header.split(',')

if columns.length < 3
  puts "Some serious problems input file, e
  return nil
end

options = Hash.new
columns.each { |column| options[column] = co

dataRequester = FormDataRequester.new(GS_USE

dataRequester.add_select_group("time", "Tim

dataRequester.request_data

timeColumn = dataRequester.get_value("time"
eventColumn = dataRequester.get_value("event
groupColumn = dataRequester.get_value("group")

puts timeColumn
```

Checkout from Experiment Repository

Select Resource

Select a resource which will be checked out as project.

AsyncSubtyping 746

DynamicKaplanMeier 735

EBI_phosphatase 1707

EchoExperiment 1697

eksperymnt1prob2 923

Experiment with Space 735

Geno2drs 875

Gnuplot_animated 851

GnuPlotExperiment 877

HeatFlow 832

HIVSubtype 744

InteractiveInterpreter 853

InterProScan 735

KaplanMeier 735

?

< Back

Next >

Finish

Cancel

Problems Tasks Ruby Documentation Console Outline Resource Browser

ViroLab Grid Resource Registry RC

ViroLab GRR

cyfronet

gridspace

demo

gem

lcg

pocket

weka

EarlyFolding

EchoService

GnuPlot

src.DynamicKaplanMeier.rb - DynamicKaplanMeier

08:39

4 of Mar 2009

with
s

tory

orm

**Certificate Info**

Type	X.509
Version	v3
Subject DN	CN=Mariusz Sterzel, O=Cyfronet, O=GRID, C=PL
Issuer DN	CN=Polish Grid CA, O=GRID, C=PL
Valid from	05.01.2009
Valid until	04.02.2010

Proxy Info**Applications****Extra Widgets****Execution Summary**



Certificate Info

Proxy Info

User proxy file missing

[Configure ...](#)

Applications

Extra Widgets

Execution Summary

Proxy Upload

Proxy Upload Form

Proxy location

[Browse...](#)

[Upload proxy](#)

VOMS Proxy Initialization

User credentials

Certificate file


[Browse...](#)

Key file

[Browse...](#)

[Upload files](#)

Current credential files

 No files

Proxy Generation

Certificate password

Virtual Organization

[Generate proxy](#)



Certificate Info

Proxy Info

User proxy file missing

[Configure...](#)

Applications

Extra Widgets

Execution Summary

Proxy Upload

Proxy Upload Form

Proxy location

[Browse...](#)

[Upload proxy](#)

VOMS Proxy Initialization

User credentials

Certificate file


[Browse...](#)


Key file


[Browse...](#)

[Upload files](#)

Current credential files

 Current credential files

 usercert.pem

 userkey.pem

Proxy Generation

Certificate password

Virtual Organization

voce

[Generate proxy](#)



Certificate Info

Proxy Info

User name	/C=PL/O=GRID /O=Cyfronet /CN=Mariusz Sterzel
Hours left	11
Issuer	/C=PL/O=GRID /O=Cyfronet /CN=voms.cyf-kr.edu.pl
VO Name(s)	/gaussian /Role=NULL /Capability=NULL
Configure ...	

Applications

Extra Widgets

Execution Summary

Proxy Upload

Proxy Upload Form

Proxy location [Browse...](#)

[Upload proxy](#)

VOMS Proxy Initialization

User credentials

Certificate file [Browse...](#)

Key file [Browse...](#)

[Upload files](#)

Current credential files

- Current credential files
 - usercontent.pem
 - userkey.pem
 - x509_proxy

Proxy Generation

Certificate password

Virtual Organization [gaussian](#)

[Generate proxy](#)

Gaussian

Extra Widgets

File Upload

Ch

Gaussian Management

Gaussian setup

% Section

Route Section

Title Section

Charge

Multiplicity

Molecule Specification

LFC file location

%Mem=512Mb
%chk=preopt.chk

#p b3lyp/cc-pVTZ symmetry=veryloose opt(tight) freq

b-carotene-15 geopt + freq

0

1

C	3.52070	-2.11940	-0.41550
C	4.86020	-2.11940	-0.41550
H	2.95120	-1.17530	-0.41550

Read

SPECIFICATION

%SECTION

%chk=myjob.chk
 %Mem=512 MB
 %Nproc=1

KEYWORDS

#p rbp86/aug-cc-pVTZ

TITLE

My very important computation

CHARGE / MULTIPL.

0, 1

GEOMETRY

shift=0
each=1
each=1
shift=0

ADD. PARAMETERS

SAVE

Charge / Multiplicity

Geometry Selection

Additional Parameters

Section

Job Type

Method

Basis Set

Job Title

BASIS SET

BASIS SET:

cc-pVTZ

▼

▼

▼

▼

(

▼

,

▼

)

☒ aug

CUSTOM:

< PREV

NEXT >

SPECIFICATION

%SECTION

%chk=myjob.chk
%Mem=512 MB
%Nproc=1

▶

KEYWORDS

#p rbp86/aug-cc-pVTZ

▶

TITLE

▶

CHARGE / MULTIPL.

0,1

▶

GEOMETRY

shift=0
each=1
each=1
shift=0

▶

ADD. PARAMETERS

▶

SAVE

Charge / Multiplicity

Geometry Selection

Additional Parameters

Section

Job Type

Method

Basis Set

Job Title

METHOD

METHOD:

Ground State

DFT...

BP86

☒ Restri
 ☐ Unres

STATES

Default

NUMBER OF STATES

NStates=

STATES OF INTEREST

Root=

GUESS METHOD:

Default

Checkpoint File

%chk=

☐ Use Quadratically Convergent SCF
 ☐ Ignore Symmetry
 ☒ Additional Print

< PREV

NEXT >

RI



Certificate Info

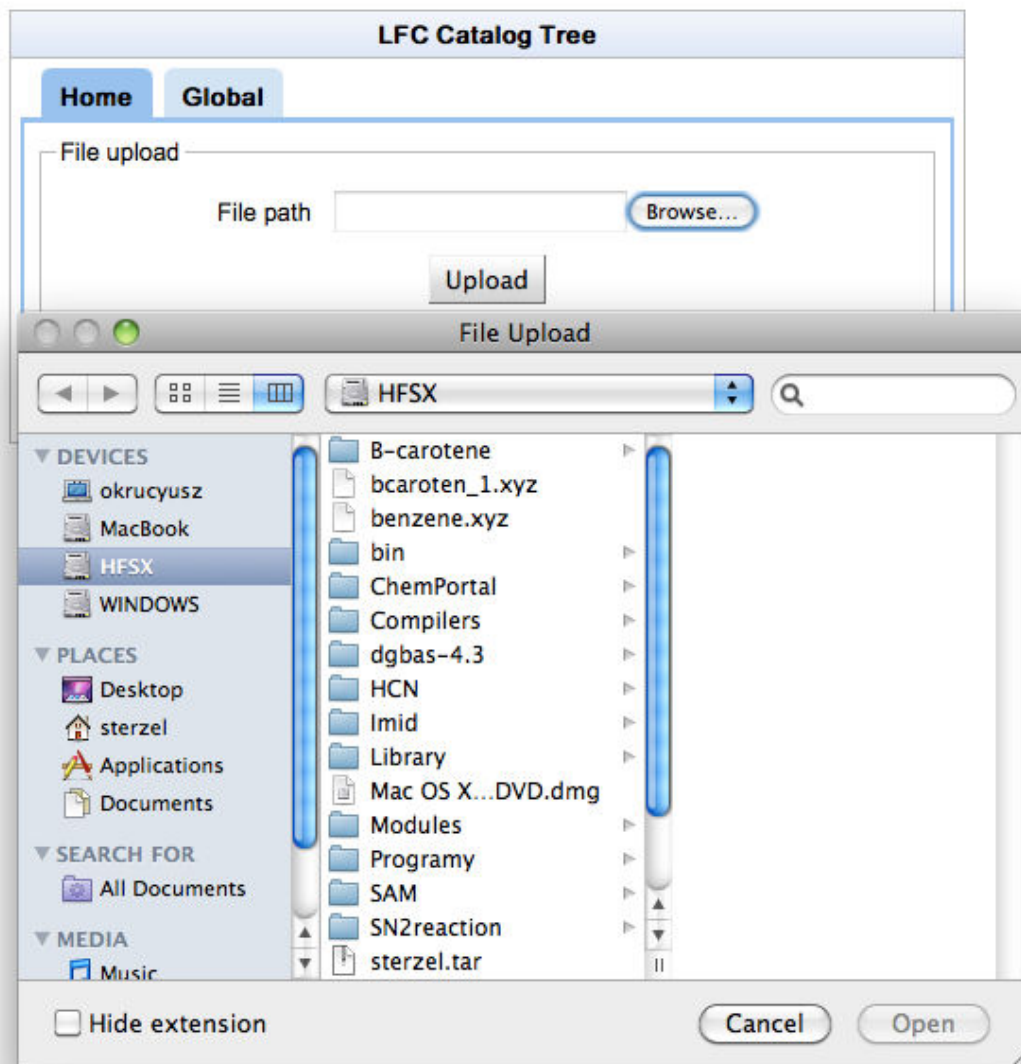
Proxy Info

Applications

Extra Widgets

LFC Tree

Execution Summary





Powered by GridSpace

Certificate Info

Proxy Info

Applications

Extra Widgets

Execution

Active sessions

Show

Execution Details

Experiment name and version	Gaussian3 v0.1	Show log
Experiment status	initialized	
LFC working directory	\$LFC_CHEMPO_USER_HOME/chempo-experiments /9070-1252678784869-016-93172	
External jobs	0	

Jobs details

Show jobs

Job ID	https://lb.grid.cyf-kr.edu.pl:9000/qpgtfPvDti0jb9MHelfyzQ
Job Status	SCHEDULED

Close

Current work:

... fixing, polishing, experiments development...

Future plans:

- API for myproxy
- Integration with visualization tools (like Jmol)
- Development of small widgets helping user to chase interesting information from output files
- Integration with GAMESS and NAMD

Questions?