

Lifeportal - HPC enabled Galaxy instance at the University of Oslo

<https://lifeportal.uio.no>

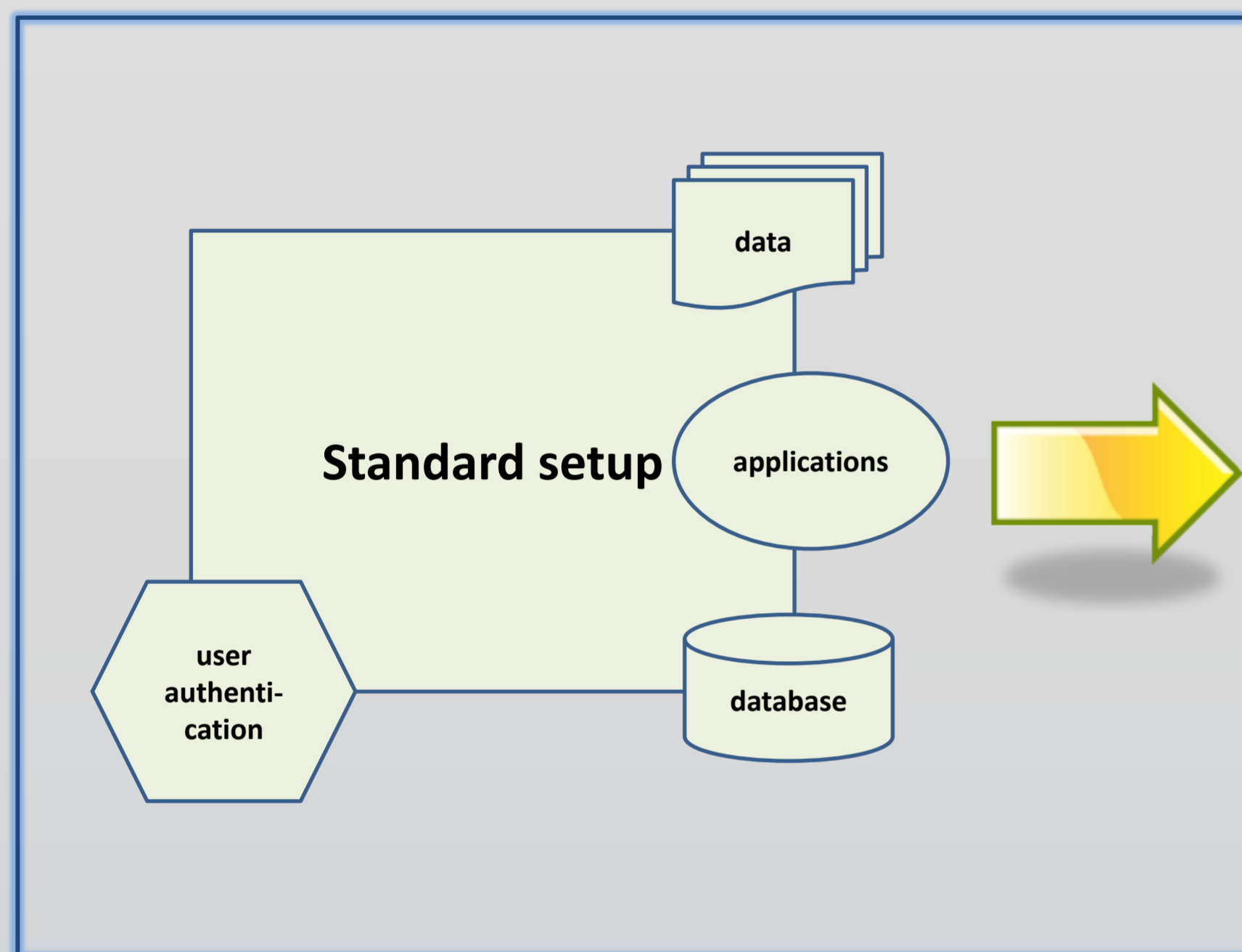
Nikolay Vazov, Katerina Michalickova, Gard Thomassen, Hans A. Eide
University Center for Information Technology, University of Oslo

One of the main goals of the high performance computing services at the University of Oslo is to make the complex HPC resources accessible to a wide audience with a varied degree of experience.

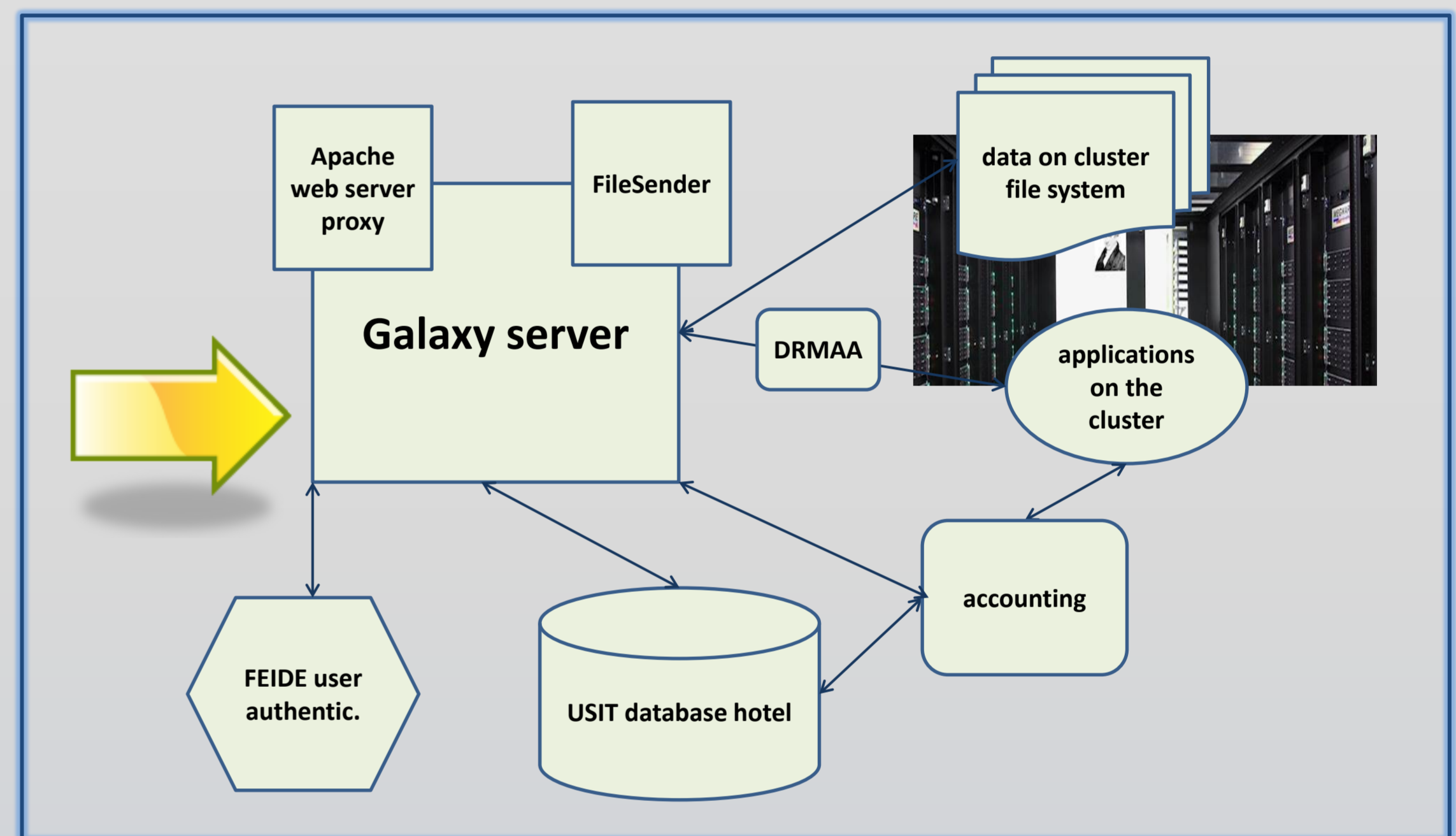
The Lifeportal enables users to utilize the HPC resources through a web interface making them accessible not only to researchers traditionally close to computing disciplines but also to researchers from domains with emerging needs for computational resources (e.g. medical research or linguistics). The Lifeportal is currently geared towards biomedical research with a special emphasis on the next generation sequencing data processing while a text mining instance is being finalized.

The front end of the Lifeportal is based on Galaxy - web-based platform for data intensive research. Apart from providing an easy to use web interface to a computing backend, Galaxy enables collaborative and reproducible research by controlled data and workflow sharing. Integration of new application into the Lifeportal is greatly enhanced by an existing public repository of application wrappers (toolshed.g2.bx.psu.edu/).

In our poster, we describe additions and modifications that lead from a standard Galaxy distribution to the production Lifeportal server.



- ### Lifeportal features
- Cluster integration
 - Web proxy-server Apache and SSL
 - Databases on a separate DB host (SSL connection)
 - Remote authentication and authorization (FEIDE - Norwegian Academic IDP)
 - Accounting with integrated project management interface:
 - Users are organized hierarchically (project managers and users)
 - GOLD allocation manager accounts for the resources management (users/projects/cpu hrs)
 - Big file upload (Filesender) – up to 250 GB per file
 - Reporting module generating PDF reports on the fly



- ### Accounting
- Only one Galaxy project on the cluster
 - Sub-project management on Galaxy level (GOLD manager)
 - Galaxy-integrated project application, project management and project reporting interface
 - Types of users:
 - Visiting user – instant access to limited resources from the portal's pool
 - Small project access – locally granted access to resources from the portal's pool
 - Large projects – reviewed projects with access to resources outside the portal's pool

ID	Name	Description	Projects	Users	Machines	Amount
11	staff	Auto-generated	staff	ANY	True	0
12	lab	Auto-generated	lab	ANY	True	0
17	ip8	account for ip8 project	ip8	ANY	True	0
19	ip9	account for ip9 project	ip9	ANY	True	0
35	hede007@hede.no_gx_default	gx_default	hede007@hede.no	ANY	True	200
36	hede008@hede.no_gx_default	gx_default	hede008@hede.no	ANY	True	200
17	hede009@hede.no_gx_default	gx_default	hede009@hede.no	ANY	True	300

Lifeportal project application form

Please fill in the required information

Personal information about the project leader

Project responsible, e.g. John Doe. (permitted chars: capital/small letters and blanks):

Job title/position (permitted chars: capital/small letters, digits and blanks):

E-mail address:

Phone no (please, use digits only, e.g. 0047XXXX for Norway):

Cell-phone number (please, use digits only, e.g. 0047XXXX for Norway):

Institution (Faculty, Department) (permitted chars: capital/small letters and blanks):

Country (permitted chars: capital/small letters and blanks):

Project information

Project name (permitted chars: capital/small letters and digits):

CPU hours (specify how many CPU hours you need for the project)(permitted chars: digits):

Preferred applications* (Click on the box to display the apps):

Project description (permitted chars: all):

Start date*:

End date:

I declare that the project does not contain sensitive data:

I have read the Lifeportal requirements and accepted them*:

Note: Fields marked * must be filled out.

Send application form Cancel

Name	Users	Projects	Start Date	End Date
nikolai.vazov@gmail.com	MEMBERS	toto proj	2013-10-24	2014-11-24
nikolai.vazov@gmail.com	MEMBERS	ip8 desc	2013-12-12	2013-12-12
nikolai@uio.no	MEMBERS	toto2	2013-11-13	2014-11-23
nikolai@uio.no	MEMBERS	testnoapprove1	2013-11-01	2020-05-23
nikolai@uio.no	MEMBERS	testnoapprove2	2013-11-01	2014-11-04
nikolai@uio.no	MEMBERS	testnoapprove3	2013-11-01	2014-11-04
nikolai@uio.no	MEMBERS	testnoapprove4	2013-11-01	2014-11-04
nikolai@uio.no	MEMBERS	testnoapprove5	2013-11-01	2014-11-04
support@feide.no	MEMBERS	test04		
support@feide.no	MEMBERS	reportproject	2014-03-31	2014-09-30
support@feide.no	MEMBERS	CLOTU project		
vazov@abv.bg	MEMBERS	ip5	2013-09-01	2015-09-15
vazov@abv.bg	MEMBERS	ip7 aa	infinity	infinity
vazov@abv.bg	MEMBERS	ip9 desc	2013-09-30	2016-09-05

- ### Cluster job submission
- Building blocks
 - SLURM job manager <https://computing.llnl.gov/linux/slurm/>
 - SLURM DRMAA (<http://apps.man.poznan.pl/trac/slurm-drmaa>)
 - DRMAA python egg
 - Cluster file system exported to the Lifeportal server
 - Galaxy file edits
 - drmaa.py, mapper.py, tool_runner.py
 - User-defined job parameters – tool_form.mako

Lifeportal

Abel handshaker (version 1.0.0)

JOBS PARAMETERS

Projects/Accounts

Number of tasks per node

Number of tasks per node

Walltime (job duration)

Memory per CPU

Available Applications:

- Abel handshaker
- BEAST
- BLAST
- Gaussian
- Migrate
- MrBayes
- RAxML
- PAUP
- PhyML
- MrModeltest
- PAML
- PhyloBayes
- Gaussian
- Autodock
- Autodock screen

- ### Applications
- Currently about 350 applications available
 - Applications reside and are executed on the HPC resource
 - PATH variable managed by environment modules package
 - All Galaxy internal scripts exported to cluster file system (galaxy-dist/lib source tree)
 - Locally developed wrappers - MrBayes, BEAST, SNAPP, Structure, RAxML, PAUP, PhyML, Migrate, MrModeltest, PAML, PhyloBayes, Gaussian, Autodock, Autodock screen
 - Filesender integration for large files upload (https://www.assembla.com/spaces/file_sender)