

# High-level APIs for managing computations on the HPC systems

**Marek Kasztelnik, Tomasz Gubała, Piotr Nowakowski, Jan Meizner, Piotr Połec, Maciej Malawski, Marian Bubak**

Academic Computer Centre Cyfronet AGH University of Science and Technology, Kraków, Poland, <https://cyfronet.pl>  
Sano Centre for Computational Medicine, Kraków, Poland, <https://sano.science>



# Running computations on the Prometheus HPC - typical routine approach

To run a computation on Prometheus usually following steps are performed:

- Copy all inputs (e.g. by using *sftp* - plgrid username and password is required)
- Copy computation code (e.g. by using *sftp* - plgrid username and password are required)
- Create slurm starting script and copy it to Prometheus
- Use ***sbatch*** to start the calculation
- Monitor started job (e.g. by using ***pro-jobs***)
- To retrieve job statistics, you can use ***sacct***
- During execution (and after computation is finished) you can monitor *stdout* and *stderr* by displaying the generated *stdout* and *stderr* files (e.g. *tail -f std\**)
- After the calculation is finished you can download results (once again e.g. by using *sftp* - PLGrid username and password are required)

The problem:

- Lots of manual work needs to be done (copy files, log in to the cluster, run many commands to start and monitor execution)
- It is hard (or even impossible) to integrate this kind of workflow with third-party application (e.g. workflow/pipeline management tool) because nobody (:-)) will paste PLGrid username/password to a third-party application

# Problem - how to integrate scientific portal with Prometheus?

Scientific gateway (web-based application)

???



# Solution (??) - user rights delegation by using grid proxy certificate

Scientific gateway (web-based application)

user proxy certificate

GSI-SSH  
GridFTP



User proxy certificate - short-lived certificate signed by long-lived certificate

GSI-SSH - open ssl connection to remote server where user proxy certificate is used to authenticate and delegate user rights

GridFTP - open FTP connection to remote server where user proxy certificate is used to authenticate and delegate user rights



# Solution (??) - user rights delegation by using user proxy certificate

Scientific gateway (web-based application)

user proxy certificate

GSI-SSH  
GridFTP



How to get user proxy certificate?

Option1: Log in to Prometheus cluster and run **grid-proxy-init**

```
(base) [prometheus][plgkasztelnik@login02 ~]$ grid-proxy-init
Your identity: /C=PL/O=PL-Grid/O=Uzytkownik/O=PL-Grid/CN=Marek Kasztelnik/CN=plgkasztelnik
Enter GRID pass phrase for this identity:
Creating proxy ..... Done
Your proxy is valid until: Tue Apr  5 00:27:50 2022
(base) [prometheus][plgkasztelnik@login02 ~]$ ls -l /tmp/x509up_u100630
-rw----- 1 plgkasztelnik plgrid 4864 Apr  4 12:27 /tmp/x509up_u100630
(base) [prometheus][plgkasztelnik@login02 ~]$
```

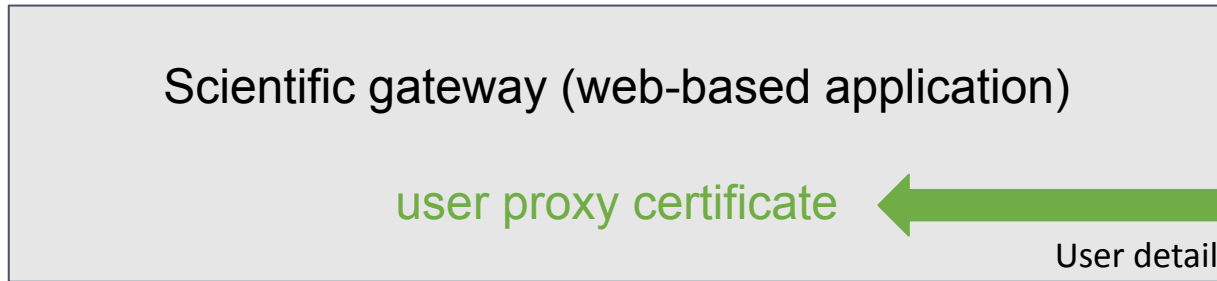
and upload the generated proxy to the scientific gateway.

This solution has the following drawbacks:

- The proxy is only valid for a short period of time and needs to be refreshed frequently
- We wanted to avoid the command line, didn't we?



# Solution (??) - user rights delegation by using user proxy certificate



Login

User details and  
user proxy certificate



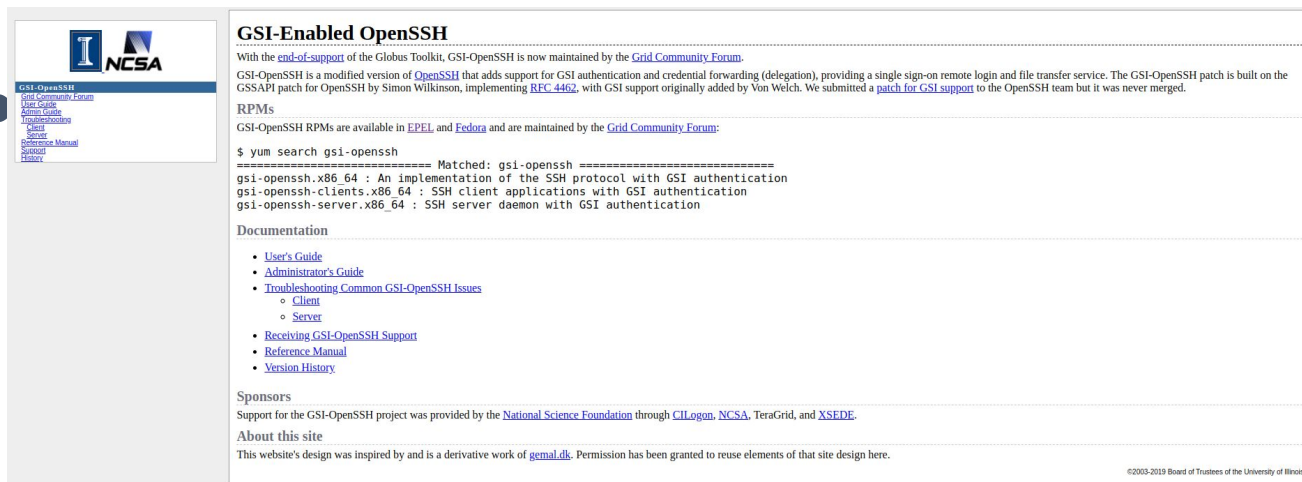
GSI-SSH  
GridFTP



## Option2: Use PLGrid IDP (identity provider)



# GSI SSH, GridFTP - how to install and use it?



The screenshot shows the website for GSI-Enabled OpenSSH. It features the NCSA logo and a navigation menu with links to Home, User's Guide, Administrator's Guide, Troubleshooting, Client, Server, Reference Manual, and Version History. The main content area is titled "GSI-Enabled OpenSSH" and includes a paragraph about the project's maintenance by the Grid Community Forum. Below this, it lists RPMs available in EPEL and Fedora. A terminal snippet shows the output of a yum search for gsi-openssh. The page also contains sections for documentation (User's Guide, Administrator's Guide, Troubleshooting, etc.) and sponsors (National Science Foundation).

## Problems:

- Not easy to install
- C based implementation
- Java implementation is outdated and not maintained anymore
- No bindings for modern programming languages

Table of Contents

- Building and installing
  - Building and installing GridFTP
  - Building and installing a static GridFTP server
  - Enabling Threading in GridFTP
- Configuring GridFTP
  - GridFTP server configuration overview
  - Typical configuration
  - Firewall requirements
  - Configuring Security for GridFTP
  - globus-gridftp-server quickstart
  - Enabling File Sharing with Globus Sharing Service
- Key Admin Settings and Tuning Recommendations
  - Concurrent Instances
  - Disk Block Size

## GCT 6.2 GridFTP : System Administrator's Guide

**NOTE** The Grid Community Toolkit documentation was taken from the Globus Toolkit 6.0 documentation. As a result, there may be inaccuracies and outdated information. Please report any problems to the Grid Community Forums as [GitHub issues](#).

[GCT](#) → [GridFTP](#) → GCT 6.2 GridFTP : System Administrator's Guide

### Introduction

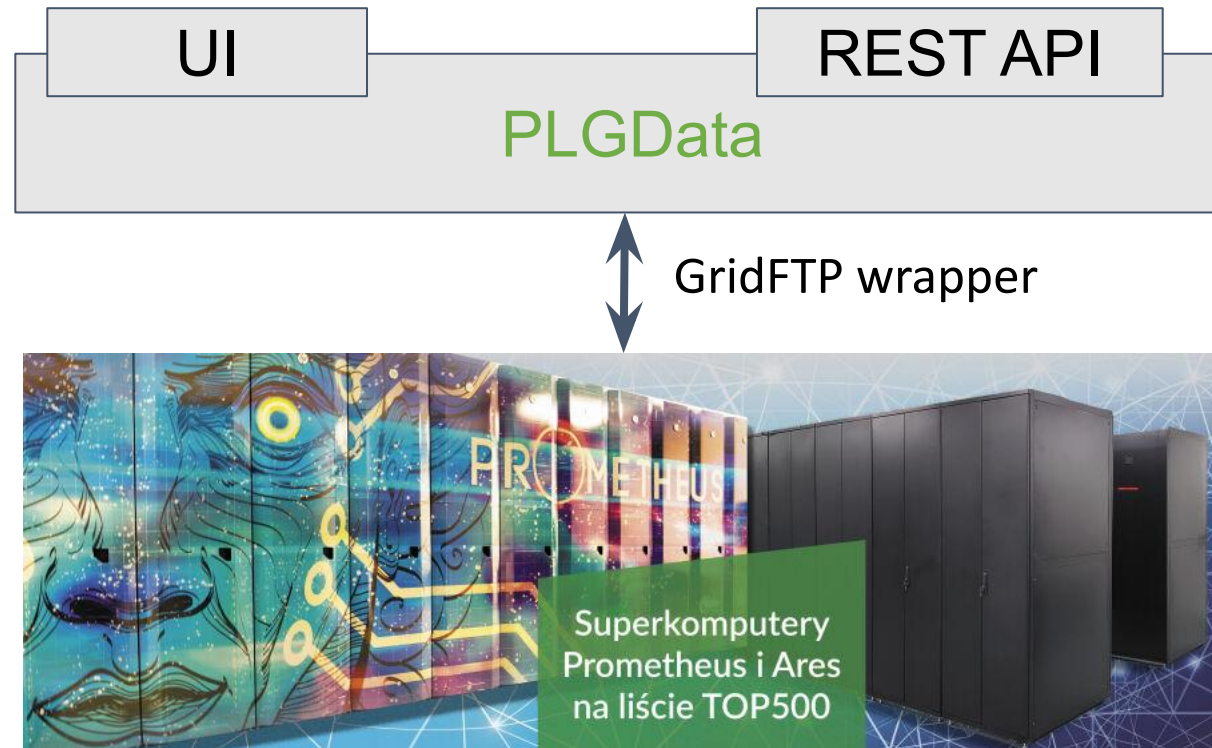
This guide contains advanced configuration information for system administrators working with GridFTP. It provides references to information on procedures typically performed by system administrators, including installation, configuring, deploying, and testing the installation. This guide should help you configure and run the GridFTP server in some standard configurations.

This information is in addition to the basic Grid Community Toolkit prerequisite, overview, installation, security configuration instructions in the [Installing GCT 6.2](#). Read through this guide before continuing!



# PLGData to the rescue - files management

- PLGrid Data (<https://data.plgrid.pl>) is a web interface and set of REST APIs for managing files stored on Prometheus.
- It is integrated with the PLGrid security system
- If you have a PLGrid account and access to Prometheus turned on, and have generated a SimpleCA certificate for access delegation, you are able to use this intuitive tool for file management.





# PLGData to the rescue - how it looks like

PLG-Data Folder shortcuts Zeus Folder shortcuts Prometheus Sign out EN PL

LISTING FOLDER: [NET/](#) / [PEOPLE/](#) / [PLGKASZTELNIK/](#)

Upload files

New directory

Total files, including 35 hidden. [Show them.](#)

Rights	Size	Modification date	Name	Type
drwxr-----	2 KB	Sep 15 10:24	<a href="#">Desktop</a>	Dir <span>Delete</span>
drwxr-xr-x	8 KB	Nov 9 12:05	<a href="#">ansys</a>	Dir <span>Delete</span>
-rwxr--r--	220 B	Oct 29 17:05	<a href="#">cleanup-ansys-p2277-20032.sh</a>	File <span>Delete</span>
drwx-----	2 KB	Jan 20 15:10	<a href="#">data.pre</a>	Dir <span>Delete</span>
-rw-r--r--	19.4 KB	Sep 30 12:22	<a href="#">error.err</a>	File <span>Delete</span>
-rw-r--r--	149 B	Oct 29 17:05	<a href="#">file0.log</a>	File <span>Delete</span>
-rw-r--r--	0 B	Oct 29 17:05	<a href="#">file0.page</a>	File <span>Delete</span>
-rw-r--r--	79 B	Oct 29 17:05	<a href="#">file1.err</a>	File <span>Delete</span>

PLG-Data Folder shortcuts Zeus Folder shortcuts Prometheus Sign out EN PL

## API Documentation

**Note: all following operations require so-called user proxy certificate.** More information related to that certificate is provided in the end of this documentation. All following code samples assume, that the proxy certificate resides in `grid_proxy` file in the current catalogue. That certificate may be passed as a parameter of a call named `proxy` or inside a header named `PROXY`. Both these options are shown in examples below. In the second case, however, the proxy certificate needs to be properly encoded.

All described remote operations on the intermediate PLGData server are performed using the HTTPS protocol, by calling a specified HTTP verb (e.g. GET or POST) at a given URL, accompanied by required parameters.

Paths to folders and files presented in examples below refer to resources stored on disks of the Zeus supercomputer. If you'd like to use API to manage files on the Prometheus supercomputer, you need to alter URLs in the examples below, adding the name `prometheus` right after the operation name. For instance, listing contents of a folder located on the Prometheus disk would require passing a URL of the following structure:

`https://data.plgrid.pl/list/prometheus/[folder_path]`, while uploading a file to a folder on Prometheus would require this: `https://data.plgrid.pl/upload/prometheus/[target_folder_path]`. Prometheus URLs for other operations would have analogous structure.

### Listing a folder

**GET:**

`https://data.plgrid.pl/list/[folder_path]`

As a call parameter one should pass the content of the proxy certificate in parameter `proxy`. Calling this operation returns a JSON document that lists content of the specified folder including subfolders and hidden files. The resulting

```
~ % curl -X POST https://data.plgrid.pl/upload/people/plguserlogin/zzzz  
-F proxy="`cat grid_proxy`" -F "file=@graph.png"
```

# Rimrock to the rescue - jobs managements

- Rimrock (<https://rimrock.plgrid.pl>) delivers a REST APIs to talk to integrated HPC infrastructures, including Prometheus
- All REST requests are secured by proxy certificates, which enable delegation of user identities

REST API

Rimrock



GSI-SSH wrapper



rimrock

PROCESSES JOBS TEAM PL EN

Submit new job Get information about all jobs Get job info Delete job Abort job

URL `/api/jobs`

Method **POST**

URL params *None*

Data params **JSON:**

```
{
  "host": [string (required)],
  "working_directory": [string (optional)],
  "script": [string (required)]
  "tag": [string (optional)]
}
```

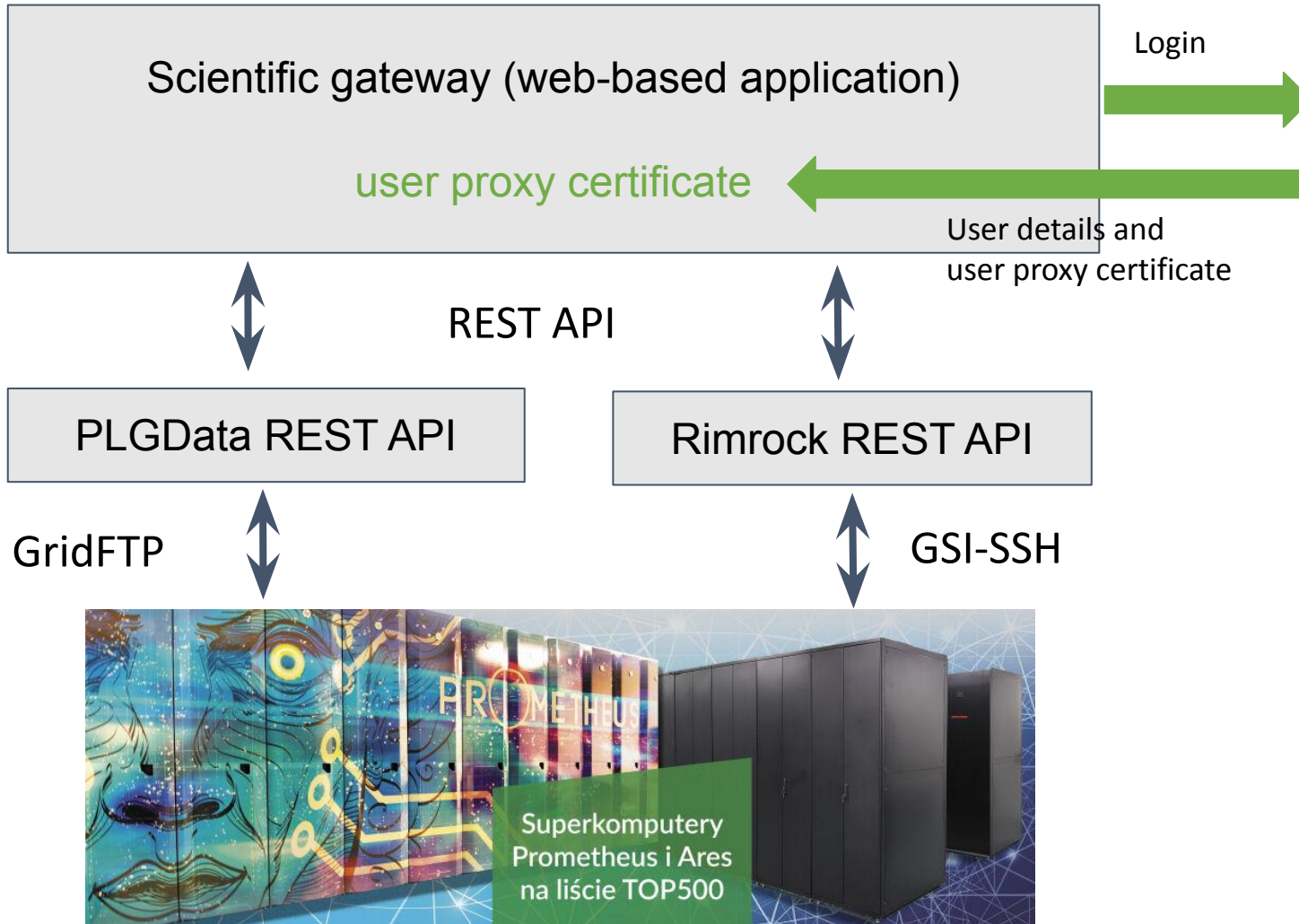
**Warning!** By default `working_directory` is set to user home directory. As a consequence when two or more jobs are started with default `working_directory` value at the same time then job script may be overwritten.

**Example:**

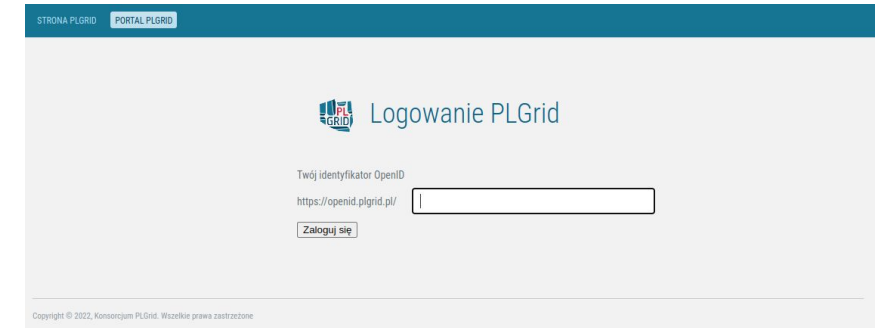
```
{
  "host": "zeus.cyfronet.pl",
  "working_directory": "/people/username/testjob",
  "script": "#!/bin/bash\nnecho hello\nexit 0"
}
```

```
curl -k -X POST --data '{"host":"pro.cyfronet.pl", "script":"#!/bin/bash\n#SBATCH -A {grantid}\necho hello\nexit 0"}' \
--header "Content-Type:application/json" --header "PROXY:$proxy" https://rimrock.plgrid.pl/api/jobs
```

# Solution !!!



## Option2: Use PLGrid IDP



## High-level service to manage data and computations in the context of a patient cohort.

- Select patient (or group of patients)
- Choose computation pipeline (series of calculation with input/output dependencies)
- Start the pipeline
- Browse the results

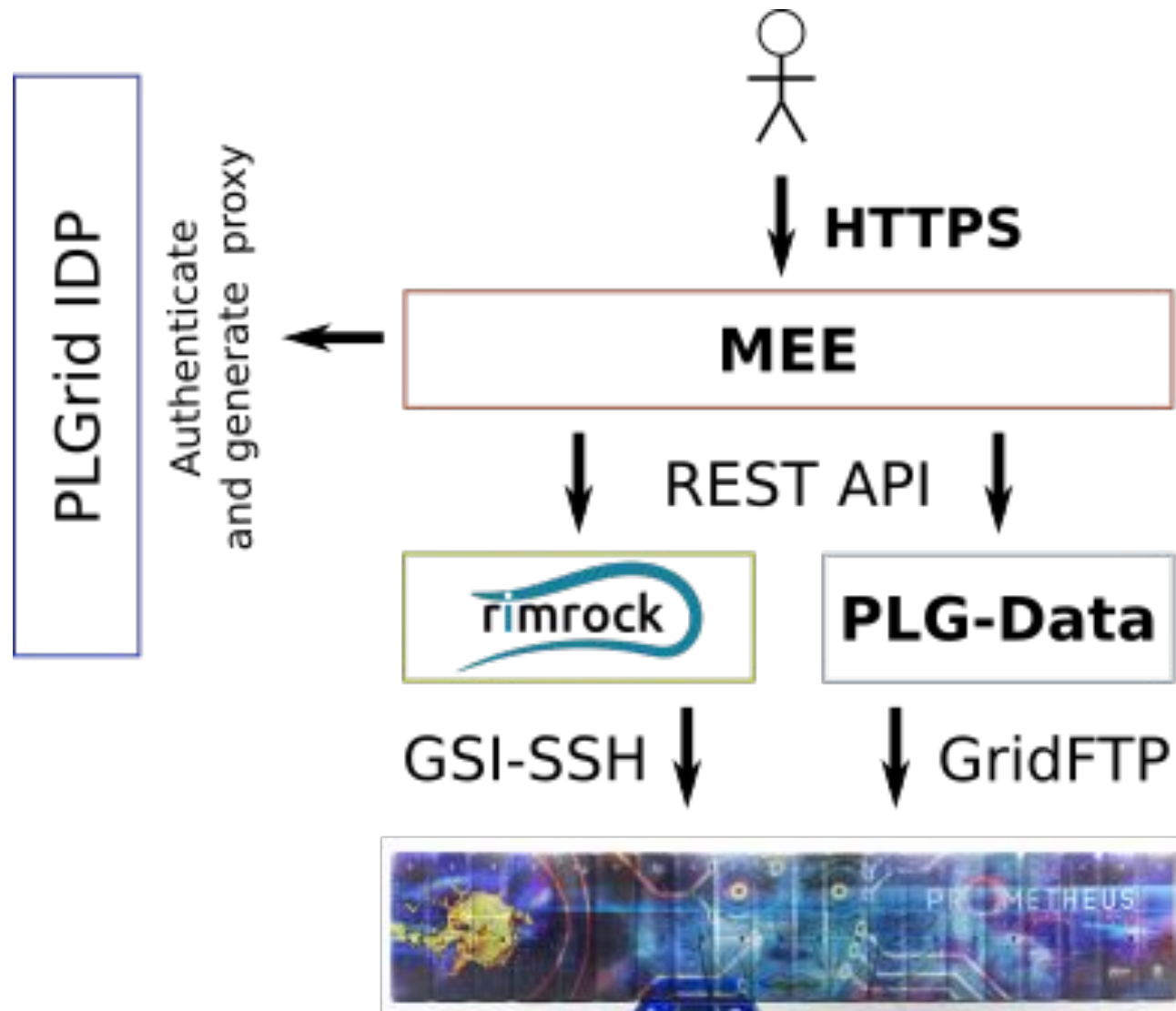
### Features:

- Integrated with Prometheus (automatic user credential delegation)
- Selecting data, starting the calculations, browsing the results in one place
- Automatic and manual pipeline execution

The screenshot displays the PRIMAGE web interface. The top navigation bar includes the PRIMAGE logo, a user profile for Marek Kasztelek, and buttons for 'Edit pipeline' and 'Remove pipeline'. The main content area is titled '/ wp5-models-test / test WP5 models (manual pipeline)'. It features a status bar indicating 'AMB finished successfully, results stored in the outputs directory'. Below this, there are sections for 'Computation details' (including start time, revision, execution time, outputs, and status) and 'Saved parameter values' (showing grant and model version). At the bottom, there are input sections for 'test pipeline inputs' and 'wp5-models-test patient inputs', both currently empty, and a table of 'test pipeline outputs' with columns for file name and data type.

File name	Data type
amb-output.json	Unspecified type
diffusion.out	Unspecified type
structural.out	Unspecified type

# Real world example -> MEE architecture



## Basic flow:

- When user logs in, a short-lived user certificate is stored in the DB
- When browsing pipeline inputs and outputs the proxy certificate is used to delegate user rights to the infrastructure by using PLGData REST API
- When calculation is started or managed, the proxy certificate is used to delegate user rights to the infrastructure by using Rimrock REST API

# MEE - running computations

1. **MEE** checks if all required inputs are present and shows an error when inputs are missing
2. Fetch model starting script template from GIT repository in the version specified by the user
3. Generate model starting script basing on the inputs provided by the user
4. Submit job to HPC queuing system (Slurm)
5. Upload required inputs to HPC
6. Monitor job execution
7. Fetch results from HPC

# MEE - running computations

1. MEE checks if all required inputs are present and shows an error when inputs are missing
2. Fetch model starting script template from GIT repository in the version specified by the user
3. Generate model starting script basing on the inputs provided by the user
4. Submit job to HPC queuing system (Slurm)
5. Upload required inputs to HPC
6. Monitor job execution
7. Fetch results from HPC

```
1 #!/bin/bash -l
2 #SBATCH -N 1
3 #SBATCH --ntasks-per-node=1
4 #SBATCH --time=00:01:00
5 #SBATCH -A {{ grant_id }}
6 #SBATCH -p plgrid-testing
7 #SBATCH --output /net/archive/groups/plggprimage/slurm_outputs/slurm-%j.out
8 #SBATCH --error /net/archive/groups/plggprimage/slurm_outputs/slurm-%j.err
9
10 # Finish with error on first command with error
11 set -e
12
13 ## Change to the directory where sbatch was called
14 cd $SCRATCHDIR
15
16 ## Clone repository and switch into selected revision
17 echo Preparing computation source code
18 -----START-----
19 {% clone_repo %}
20 -----END-----
21
22 # Prepare 1 demo numbers
23 echo -----START-----
24 {% stage_in_demo_numbers %}
25 -----END-----
26
27 echo Sorting
28 -----START-----
29 python demo-steps/1_sort.py
30 -----END-----
31
32 echo Uploading results
33 -----START-----
34 {% stage_out_steps.txt %}
35 -----END-----
36 echo Finish
```

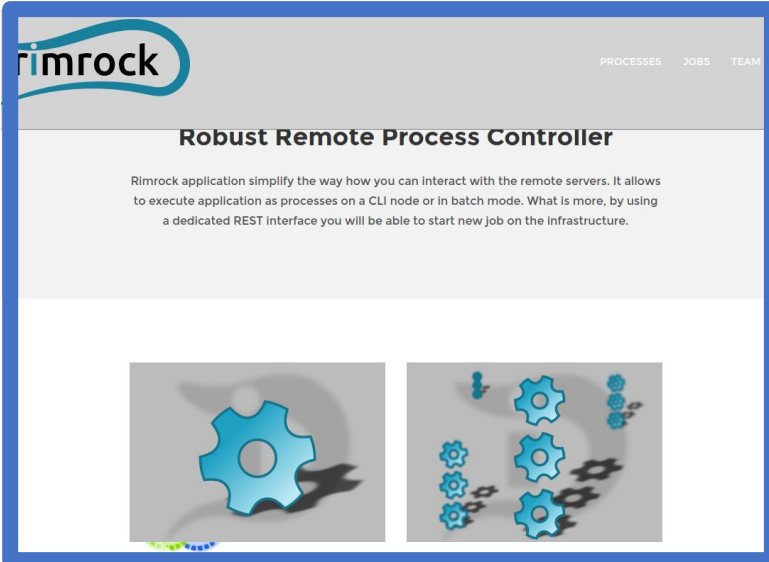
```
1 #!/bin/bash -l
2 #SBATCH -N 1
3 #SBATCH --ntasks-per-node=1
4 #SBATCH --time=00:01:00
5 #SBATCH -A plgprimage3
6 #SBATCH -p plgrid-testing
7 #SBATCH --output /net/archive/groups/plggprimage/slurm_outputs/slurm-%j.out
8 #SBATCH --error /net/archive/groups/plggprimage/slurm_outputs/slurm-%j.err
9
10 # Finish with error on first command with error
11 set -e
12
13 ## Change to the directory where sbatch was called
14 cd $SCRATCHDIR
15
16 ## Clone repository and switch into selected revision
17 echo Preparing computation source code
18 -----START-----
19 export SSH_DOWNLOAD_KEY="-----BEGIN RSA PRIVATE KEY-----
20 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
21 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
22 -----END RSA PRIVATE KEY-----"
23
24 ssh-agent bash -c '
25 ssh-add <(echo "$SSH_DOWNLOAD_KEY");
26 git clone git@gitlab.com:primageproject/mee/demo-steps
27 cd `basename primageproject/mee/demo-steps .git`
28 git reset --hard 68fc5f04f2cfb8f80d04fbb85a4a050efa225192'
29 -----END-----
30
31 echo Downloading numbers
32 -----START-----
33 cp /net/archive/groups/plggprimage/development/patients/mktest/pipelines/1/outputs/numbers.txt $SCRATCHDIR/numbers.txt
34 -----END-----
35
36 echo Sorting
37 -----START-----
38 python demo-steps/1_sort.py
39 -----END-----
40
41 echo Uploading results
42 -----START-----
43 if [ -e steps.txt ]; then
44 cp steps.txt /net/archive/groups/plggprimage/development/patients/mktest/pipelines/1/outputs/steps.txt
45 else
46 echo "Cannot stage out steps.txt because the file does not exist" 1>&2
47 exit 1
48 fi
49 -----END-----
50
51 echo Finish
```

# MEE - running computations

1. MEE checks if all required inputs are present and shows an error when inputs are missing
2. Fetch model starting script template from GIT repository in the version specified by the user
3. Generate model starting script basing on the inputs provided by the user
4. Submit job to HPC queuing system (Slurm)
5. Upload required inputs to HPC
6. Monitor job execution
7. Fetch results from HPC

```
1 #!/bin/bash -l
2 #SBATCH -N 1
3 #SBATCH --ntasks-per-node=1
4 #SBATCH --time=00:01:00
5 #SBATCH -A {{ grant_id }}
6 #SBATCH -p plorid-testing
7 #SBATCH --output /net/archive/groups/plggprimage/slurm_outputs/slurm-%j.out
8 #SBATCH --error /net/archive/groups/plggprimage/slurm_outputs/slurm-%j.err
9
10 # Finish with error on first command with error
11 set -e
12
13 ## Change to the directory where sbatch was called
14 cd $SCRATCHDIR
15
16 ## Clone repository and switch into selected revision
17 echo Preparing computation source code
18
19 {% clone_repo %}
20 echo -----END-----
21
22 {% stage_in_demo_numbers %}
23 echo -----START-----
24 {% stage_in_demo_numbers %}
25 echo -----END-----
26
27 echo Sorting
28 echo -----START-----
29 python demo-steps/1_sort.py
30 echo -----END-----
31
32 echo Uploading results
33 echo -----START-----
34 {% stage_out_steps.txt %}
35 echo -----END-----
36 echo Finish
```

```
1 #!/bin/bash -l
2 #SBATCH -N 1
3 #SBATCH --ntasks-per-node=1
4 #SBATCH --time=00:01:00
5 #SBATCH -A plqprimage3
6 #SBATCH -p plorid-testing
7 #SBATCH --output /net/archive/groups/plggprimage/slurm_outputs/slurm-%j.out
8 #SBATCH --error /net/archive/groups/plggprimage/slurm_outputs/slurm-%j.err
9
10 # Finish with error on first command with error
11 set -e
12
13 ## Change to the directory where sbatch was called
14 cd $SCRATCHDIR
15
16 ## Clone repository and switch into selected revision
17 echo Preparing computation source code
18 echo -----START-----
19 export SSH_DOWNLOAD_KEY="-----BEGIN RSA PRIVATE KEY-----
20 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
21 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
22 -----END RSA PRIVATE KEY-----"
23
24 ssh-agent bash -c '
25 ssh-add <(echo "$SSH_DOWNLOAD_KEY");
26 git clone git@gitlab.com:primageproject/mee/demo-steps
27 cd `basename primageproject/mee/demo-steps .git`
28 git reset --hard 68fc5f04f2cfb8f80d04fbb85a4a050efa225192'
29 echo -----END-----
30
31 echo Downloading numbers
32 echo -----START-----
33 cp /net/archive/groups/plggprimage/development/patients/mktest/pipelines/1/outputs/numbers.txt $SCRATCHDIR/numbers.txt
34 echo -----END-----
35
36 echo Sorting
37 echo -----START-----
38 python demo-steps/1_sort.py
39 echo -----END-----
40
41 echo Uploading results
42 echo -----START-----
43 if [ -e steps.txt ]; then
44 cp steps.txt /net/archive/groups/plggprimage/development/patients/mktest/pipelines/1/outputs/steps.txt
45 else
46 echo "Cannot stage out steps.txt because the file does not exist" 1>&2
47 exit 1
48 fi
49
50 echo -----END-----
```





# MEE - running computations

1. MEE checks if all required inputs are present and shows an error when inputs are missing
2. Fetch model starting script template from GIT repository in the version specified by the user
3. Generate model starting script basing on the inputs provided by the user
4. Submit job to HPC queuing system (Slurm)
5. Upload required inputs to HPC
6. Monitor job execution
7. Fetch results from HPC

```
1 #!/bin/bash -l
2 #SBATCH -N 1
3 #SBATCH --ntasks-per-node=1
4 #SBATCH --time=00:01:00
5 #SBATCH -A {{ grant_id }}
6 #SBATCH -p plgrid-testing
7 #SBATCH --output /net/archive/groups/plggprimage/slurm_outputs/slurm-%j.out
8 #SBATCH --error /net/archive/groups/plggprimage/slurm_outputs/slurm-%j.err
9
10 # Finish with error on first command with error
11 set -e
12
13 ## Change to the directory where sbatch was called
14 cd $SCRATCHDIR
15
16 ## Clone repository and switch into selected revision
17 echo Preparing computation source code
18 echo -----START-----
19 {% clone_repo %}
20 echo -----END-----
21
22 echo -----START-----
23 echo {% stage_in_demo_numbers %}
24 echo -----END-----
25
26 echo Sorting
27 echo -----START-----
28 echo python demo-steps/1_sort.py
29 echo -----END-----
30
31 echo Uploading results
32 echo -----START-----
33 echo {% stage_out_steps.txt %}
34 echo -----END-----
35
36 echo Finish
```

```
1 #!/bin/bash -l
2 #SBATCH -N 1
3 #SBATCH --ntasks-per-node=1
4 #SBATCH --time=00:01:00
5 #SBATCH -A plgprimage3
6 #SBATCH -p plgrid-testing
7 #SBATCH --output /net/archive/groups/plggprimage/slurm_outputs/slurm-%j.out
8 #SBATCH --error /net/archive/groups/plggprimage/slurm_outputs/slurm-%j.err
9
10 # Finish with error on first command with error
11 set -e
12
13 ## Change to the directory where sbatch was called
14 cd $SCRATCHDIR
15
16 ## Clone repository and switch into selected revision
17 echo Preparing computation source code
18 echo -----START-----
19 export SSH_DOWNLOAD_KEY="-----BEGIN RSA PRIVATE KEY-----
20 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
21 XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX
22 -----END RSA PRIVATE KEY-----"
23
24 ssh-agent bash -c '
25 ssh-add <(echo "$SSH_DOWNLOAD_KEY");
26 git clone git@gitlab.com:primageproject/mee/demo-steps
27 cd `basename primageproject/mee/demo-steps .git`
28 git reset --hard 68fc5f04f2cfb8f80d04fbb85a4a050efa225192'
29 echo -----END-----
30
31 echo Downloading numbers
32 echo -----START-----
33 cp /net/archive/groups/plggprimage/development/patients/mktest/pipelines/1/outputs/numbers.txt $SCRATCHDIR/numbers.txt
34 echo -----END-----
35
36 echo Sorting
37 echo -----START-----
38 echo python demo-steps/1_sort.py
39 echo -----END-----
40
41 echo Uploading results
42 echo -----START-----
43 if [ -e steps.txt ]; then
44 cp steps.txt /net/archive/groups/plggprimage/development/patients/mktest/pipelines/1/outputs/steps.txt
45 else
46 echo "Cannot stage out steps.txt because the file does not exist" 1>&2
47 exit 1
48 fi
49
50 echo -----END-----
```

PLG-Data A simple tool for PL-Grid files management. PL EN

Welcome!

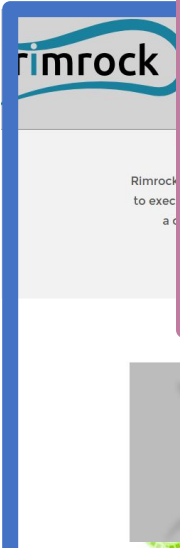
Welcome to the PLG-Data service. Using this tool you are able to access the folders and files that are available for you in the PL-Grid Infrastructure. These include both your personal resources and also those shared with you (or the teams you belong to) by collaborators.

Please use the button below to authorize yourself with the PL-Grid OpenID protocol. You will need both your PL-Grid Infrastructure login-password pair and the passphrase you used when generating the SimpleCA certificate.

Sign in with PL-Grid OpenID

Copyright © 2013-2017 Academic Computer Centre Cyfronet AGH All rights reserved. Contact us We use cookies.

Further resources: source code, API for developers.



**External applications can easily be integrated with the Prometheus cluster with:**

- **PLGrid identity provider** to log in and generate a proxy certificate for the user
- **PLGData** to manage files stored on the cluster using a modern web-based UI or REST API
- **Rimrock** to manage computations started on the cluster by using a REST API

**The usefulness of these tools has been proven by many applications integrated with the infrastructure, e.g.:**

- **Model Execution Environment** - an environment that enables computational models to be developed in a simple and organized manner and easily deployed to the available HPC infrastructure. The platform promotes managing simulations in a way that guarantees repeatability, replicability, and reproducibility
- **EPISODES Platform** - European plate induced seismicity observations and dataset platform

Rimrock and PLGData are released under an open-source license.  
Grab your version at <https://gitlab.com/cyfronet>

<http://dice.cyfronet.pl/>