



Akademickie Centrum Komputerowe
CYFRONET AGH



Marek Kasztelnik Laboratorium Metod Informatycznych w Medycynie (DICE)

Metody i narzędzia wspierające złożone
symulacje na infrastrukturach HPC

LMIM - nasz zespół

- ❑ Zespół naukowców, inżynierów oprogramowania, administratorów
- ❑ Cel: rozwój metod i narzędzi dla medycyny obliczeniowej
- ❑ Wypracowujemy metody oraz narzędzia umożliwiające uruchamianie złożonych symulacji medycznych w środowisku HPC
- ❑ Bierzemy udział w krajowych oraz międzynarodowych projektach badawczych od 2002 roku.

Ostatnich 7 projektów:

- ❑ **EDITH** (2022-2024) - foster an inclusive ecosystem for Digital Twins in healthcare in Europe
- ❑ **EuroHPC** (2021-2023) - budowa specjalistycznej infrastruktury ogólnego przeznaczenia na potrzeby obliczeń wielkoskalowych
- ❑ **PRIMAGE** (2018-2023) - creation of a Clinical Decision Support System (CDSS) for the treatment of cancer (nauroblastoma, glioma) in children.
- ❑ **Sano** (2019-2026) - Poland Europe's first Centre for simulation-driven healthcare Decision Support.
- ❑ **Gliomed** (2017-2020) - Undertaking a series of bioinformatics and biostatistics analyses on tumor data.
- ❑ **Process** (2017-2020) - Design and create prototypes for exascale data storage and processing.
- ❑ **EurValve** (2016-2019) - Creating an easy-to-use environment for the development, deployment and execution of large scale simulations, required for simulating human heart.



Marian Bubak
Team leader



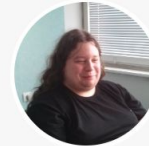
Tomasz Gubala
Researcher, programmer, entrepreneur



Marek Kasztelnik
Researcher, scientific developer



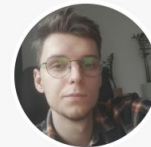
Maciej Malawski
Researcher



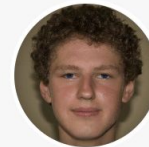
Jan Meizner
System administrator



Piotr Nowakowski
Researcher, scientific developer



Piotr Poleć
Software developer



Piotr Szczepan Wawryka
Programmer, biomedical engineer



Kanil Burkiewicz
Software developer



Spersonalizowana medycyna - aktualne problemy

- ❑ Zaawansowane modele oraz symulacje wymagają dużej mocy obliczeniowej
- ❑ Rozwiązania HPC nie są proste do użycia, szczególnie dla osób spoza dziedziny IT; potrzebna jest wiedza o: SSH, SCP, systemu kolejkowego slurm, etc.
- ❑ Rozwój modeli oraz symulacji zazwyczaj odbywa się w ramach grupy badawczej, więc potrzebne jest współdzielenie modeli, plików wejściowych oraz rezultatów
- ❑ Powinniśmy być w stanie powtórzyć obliczenia stworzone przez inną osobę, zespół badawczy
- ❑ To samo obliczenie chcielibyśmy uruchomić dla innego przypadku (np. dla innej osoby)

Jak to aktualnie zazwyczaj wygląda:

- ❑ Naukowcy zazwyczaj chcą osiągnąć wyniki jak najszybciej, skupiają się na rezultatach, często zapominając o wersjonowaniu uruchamianych modeli, rezultatów
- ❑ Kod modelu, pliki wejściowe zazwyczaj transferowane są na klaster do lokalnego katalogu użytkownika, niedostępnego dla pozostałych członków zespołu
- ❑ Kiedy obliczenia dla konkretnego przypadku zostaną zakończone, to uruchamiane są następne obliczenia, które bardzo często nadpisują rezultaty wyprodukowane przez poprzednie obliczenia i dlatego bardzo trudnym lub też często niemożliwym jest odtworzenie poprzednich rezultatów

Aby pozbyć się powyższych problemów potrzebujemy:

- ❑ Wersjonowania modeli, symulacji
- ❑ Ustrukturalizowanego sposobu przechowywania danych wejściowych dla konkretnego przypadku oraz rezultatów uruchomienia modelu, symulacji
- ❑ Możliwości odtworzenia obliczeń nie tylko przez osobę, która oryginalnie uruchamiała symulację, ale również przez innych członków zespołu badawczego

Uruchamianie obliczeń na HPC - standardowe postępowanie

- ❑ Skopiuj wszystkie niezbędne pliki wejściowe oraz kod aplikacji na węzeł dostępowy (najczęściej poprzez SCP)
- ❑ Stwórz skrypt startowy dla systemu kolejkowego (np. slurm), skopiuj go na węzeł dostępowy (SCP)
- ❑ Użyj **sbatch** aby uruchomić zadanie
- ❑ Monitoruj status zadania (**hpc-jobs**)
- ❑ Jeśli są potrzebne statystyki zadania można użyć **sacct**
- ❑ Jeśli potrzebujemy monitorować standardowe wyjście aplikacji możemy użyć **tail -f *std**
- ❑ Po zakończeniu zadania możemy pobrać pliki wyjściowe symulacji (najczęściej przez SCP)

Występujące problemy:

- ❑ Wymagane jest bardzo dużo manualnej interwencji, bardzo łatwo popełnić w którymś miejscu błąd, co może powodować błędy w uruchamianej symulacji, nadpisanie poprzednich rezultatów, etc.
- ❑ Bardzo trudno jest zintegrować taki proces z zewnętrznymi aplikacjami, np. systemami workflow
- ❑ Odtworzenie wyprodukowanych rezultatów często jest niemożliwe ze względu na brak referencji do wersji symulacji, pochodzenia plików wejściowych, etc.

Jak rozwiązujemy tego typu problemy

- ❑ Dedykowane serwisy, specyficzne dla konkretnej grupy badawczej, które wspierają konkretny tryb pracy zespołu badawczego
- ❑ Kilka przykładów:

New experiment definition

Basic information

Experiment name

Tags (separated by commas)

Data file attributes setup

Attribute name	Attribute type
Age	In
Sex	Enum
Species	Enum

Create experiment!

Integromika

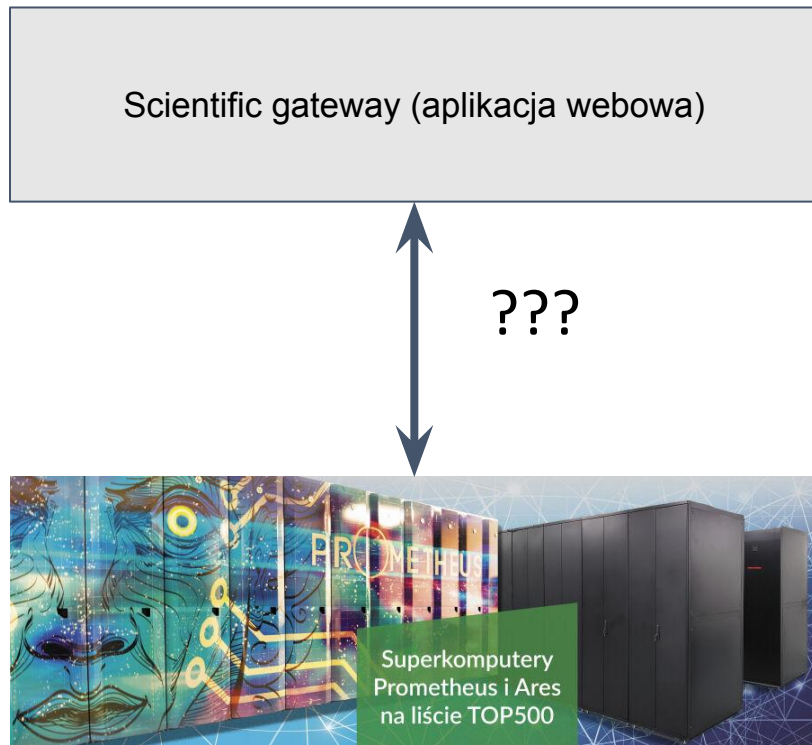
GridSpace

Meta biobank

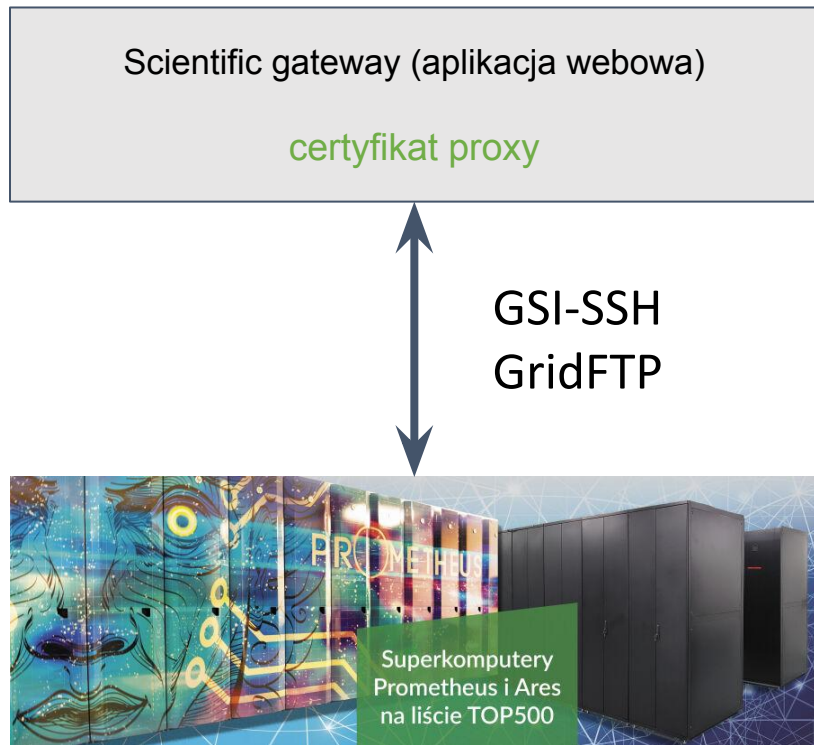
Patient code	Race	Storage temperature	Material type	Anatomical part
16	PATIENT	Caucasian	45°C to 40°C	DNA
16	PATIENT	Caucasian	45°C to 40°C	DNA
16	PATIENT	Caucasian	45°C to 40°C	DNA
16	PATIENT	Caucasian	45°C to 40°C	DNA
16	PATIENT	Caucasian	45°C to 40°C	DNA
16	PATIENT	Caucasian	45°C to 40°C	DNA
16	PATIENT	Caucasian	45°C to 40°C	DNA
16	PATIENT	Caucasian	45°C to 40°C	DNA
16	PATIENT	Caucasian	45°C to 40°C	DNA
16	PATIENT	Caucasian	45°C to 40°C	DNA
16	PATIENT	Caucasian	45°C to 40°C	DNA
16	PATIENT	Caucasian	45°C to 40°C	DNA
16	PATIENT	Caucasian	45°C to 40°C	DNA
16	PATIENT	Caucasian	45°C to 40°C	DNA
16	PATIENT	Caucasian	45°C to 40°C	DNA
16	PATIENT	Caucasian	45°C to 40°C	DNA
16	PATIENT	Caucasian	45°C to 40°C	DNA
16	PATIENT	Caucasian	45°C to 40°C	DNA
16	PATIENT	Caucasian	45°C to 40°C	DNA

Model Execution Environment

Jak zintegrować taki serwis z HPC?



Krok 1 - delegacja uprawnień użytkownika

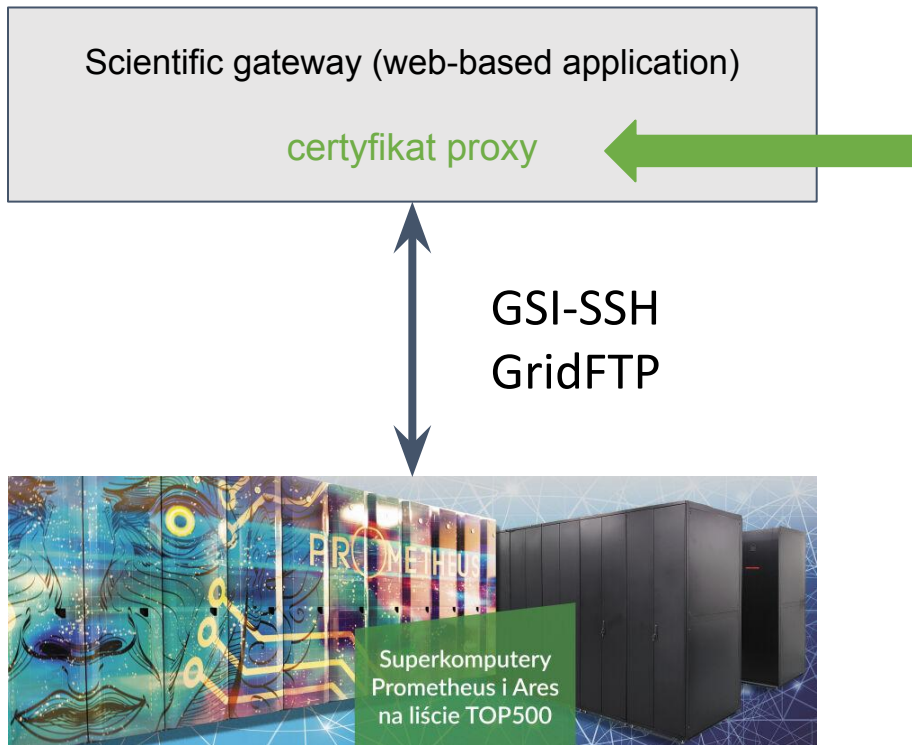


Certyfikat proxy - krótko żyjący (zazwyczaj 12h) certyfikat podpisany długo żyjącym certyfikatem

GSI-SSH - modyfikacja standardu SSL, gdzie dodana została możliwość autentykacji poprzez certyfikat proxy

GridFTP - modyfikacja standardu FTP, gdzie dodana została możliwość autentykacji poprzez certyfikat proxy

Krok 2 - skąd wziąć certyfikat proxy (1/2)?



Skąd pobrać certyfikat proxy?

Opcja nr 1:
Zalogować się na Prometheusa oraz uruchomić **grid-proxy-init**

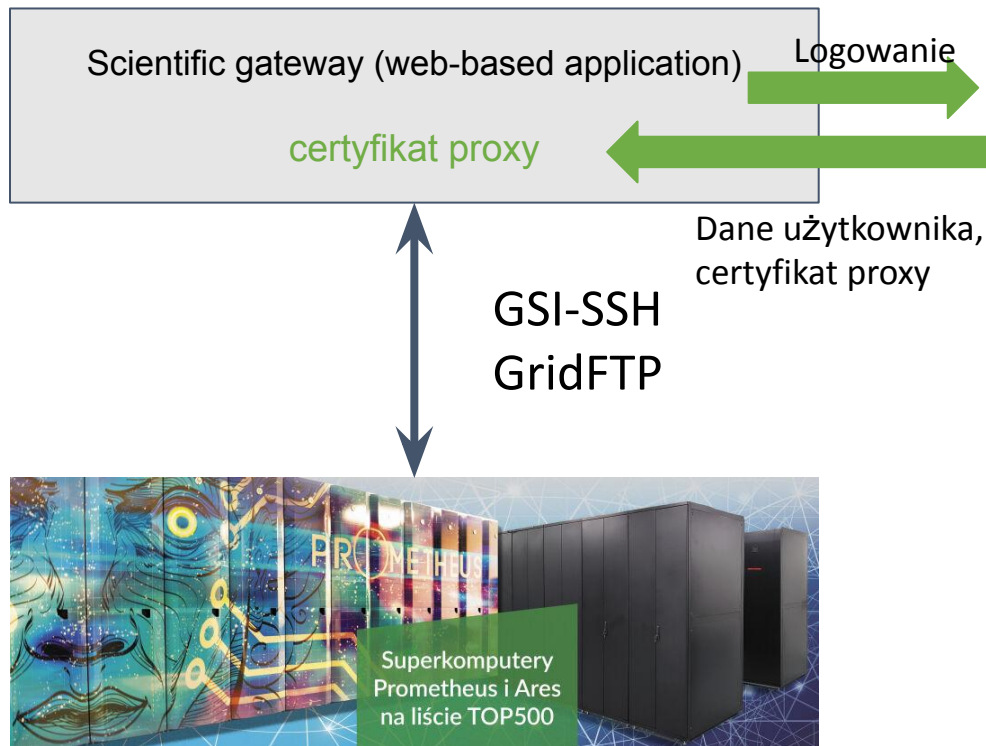
```
(base) [prometheus][plgkasztelnik@login02 ~]$ grid-proxy-init
Your identity: /C=PL/O=PL-Grid/O=Uzytkownik/O=PL-Grid/CN=Marek Kasztelnik/CN=plg
kasztelnik
Enter GRID pass phrase for this identity:
Creating proxy ..... Done
Your proxy is valid until: Tue Apr  5 00:27:50 2022
(base) [prometheus][plgkasztelnik@login02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

następnie wygenerowany certyfikat proxy możemy zaimportować do naszej aplikacji

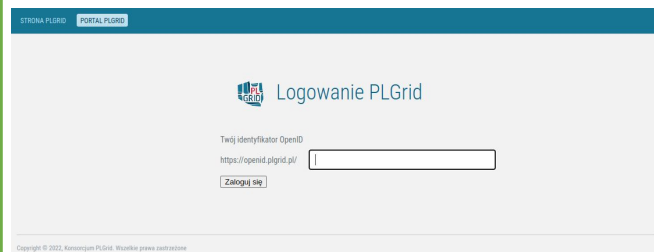
To rozwiązanie ma jednak wady:

- Ważność certyfikatu proxy jest krótka (standardowo 12h)
- Chcieliśmy uniknąć użycia linii poleceń

Krok 2 - skąd wziąć certyfikat proxy (2/2)?



Opcja 2: użyć OpenId dostarczanego przez PLGrid



Plusy:

- Nowy użytkownik nie musi zakładać konta - ma już konto PLGrid
- Brak linii poleceń, certyfikat proxy automatycznie jest "wgrzywany"

GSI SSH, GridFTP - jak z tego skorzystać



GSI-Enabled OpenSSH

With the [end-of-support](#) of the Globus Toolkit, GSI-OpenSSH is now maintained by the [Grid Community Forum](#).

GSI-OpenSSH is a modified version of [OpenSSH](#) that adds support for GSI authentication and credential forwarding (delegation), providing a single sign-on session login and file transfer service. The GSI-OpenSSH patch is built on the GSI-SSH patch for OpenSSH by Simon Wilkinson, implementing [GSI-Shell](#), with GSI support originally added by Tom White. We intend a [parallel GSI support](#) in the OpenSSH team but a year over early.

RPMs

GSI-OpenSSH RPMs are available in [EPEL](#) and [Fedora](#) and are maintained by the [Grid Community Forum](#).

```
$ yum search gsi-openssh
gsl-openssh-server.noarch : Matched: gsi-openssh
gsi-openssh-ssh_64.i386 : An implementation of the SSH protocol with GSI authentication
gsi-openssh-clients.x86_64 : SSH client applications with GSI authentication
gsi-openssh-server.x86_64 : SSH server daemon with GSI authentication
```

Documentation

- [Users Guide](#)
- [Administrator's Guide](#)
- [Troubleshooting Common GSI-OpenSSH Issues](#)
 - [Client](#)
 - [Server](#)
- [Recreate GSI-OpenSSH Support](#)
 - [Release Manual](#)
 - [Version History](#)

Sponsors

Support for the GSI-OpenSSH project was provided by the [National Science Foundation](#) through [CAREER](#), [NCSA](#), [Teragrid](#), and [NSERC](#).

About this site

This website's design was inspired by and is a derivative work of [groupl](#). Permission has been granted to reuse elements of that site design here.

©2009-2010 Board of Trustees of the University of Illinois

Problemy:

- Trudne w instalacji
- Implementacja oparta o C
- Implementacja w Java nie wspierana
- Brak wsparcia dla nowych języków programowania

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
 - glibus-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

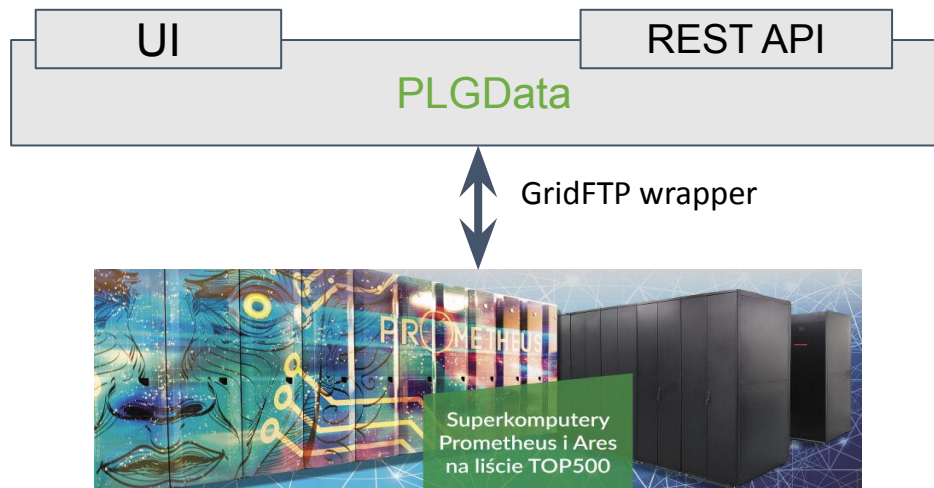
Introduction

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 - zarządzanie plikami przez GridFTP

- ❑ PLGData (<https://data.plgrid.pl>) - interfejs webowy oraz RESTowy do zarządzania plikami przechowywanymi na Prometheus, Aresie, Atenie, etc.
- ❑ Zintegrowany z systemem bezpieczeństwa PLGrid
- ❑ Jeśli jesteś użytkownikiem PLGrid oraz masz wygenerowany certyfikat SimpleCA, to możesz od razu zacząć korzystać z tego serwisu



PLGData - jak korzystać

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	Desktop	Dir
drwxr-xr-x	8 KB	Nov 9 12:05	ansys	Dir
-rwxr--r--	220 B	Oct 29 17:05	cleanup-ansys-p2277-20032.sh	File
drwx-----	2 KB	Jan 20 15:10	data_pre	Dir
-rw-r--r--	19.4 KB	Sep 30 12:22	error.err	File
-rw-r--r--	149 B	Oct 29 17:05	file0.log	File
-rw-r--r--	0 B	Oct 29 17:05	file0.page	File
-rw-r--r--	79 B	Oct 29 17:05	file1.err	File

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

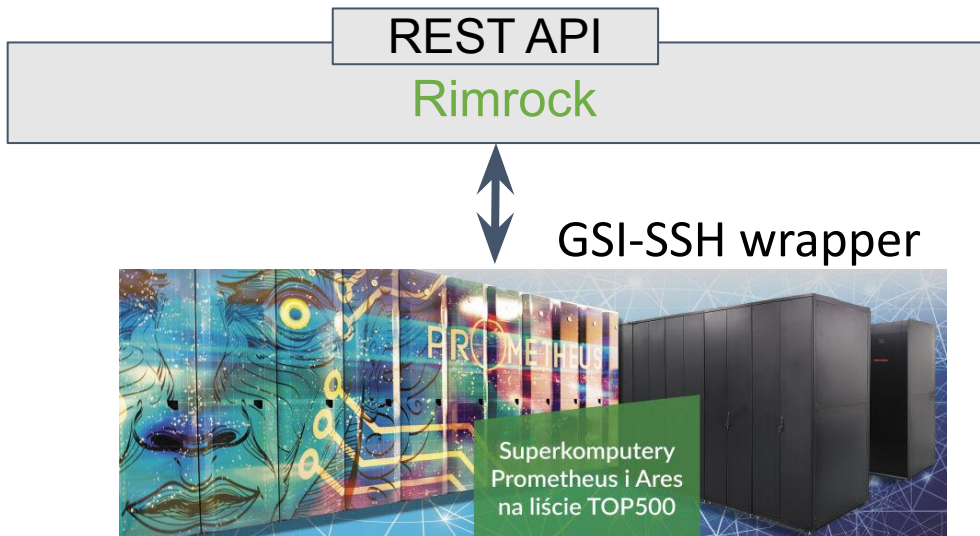
GET:
`https://data.plgrid.pl/list/[folder_path]`

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 - zarządzanie zadaniami na HPC

- ❑ Rimrock (<https://rimrock.plgrid.pl>) zbiór interfejsów RESTowych do zarządzania zadaniami na HPC, posiada wsparcie dla Prometheusa, Aresa oraz Ateny.
- ❑ Zintegrowany z systemem bezpieczeństwa PLGrid, certyfikat proxy jest użyty do delegacji uprawnień użytkownika.



The screenshot shows the Rimrock web interface. The header includes the "rimrock" logo and navigation links for "PROCESSES", "JOBS", and "TEAM". Below the header, there are links for "Submit new job", "Get information about all jobs", "Get job info", "Delete job", and "Abort job". The main content area shows a form for submitting a new job. The "URL" field is set to "/api/jobs". The "Method" is set to "POST". The "URL params" field is empty. The "Data params" field contains a JSON object:

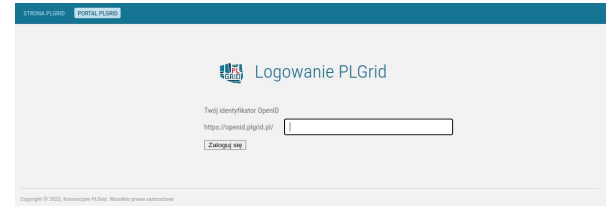
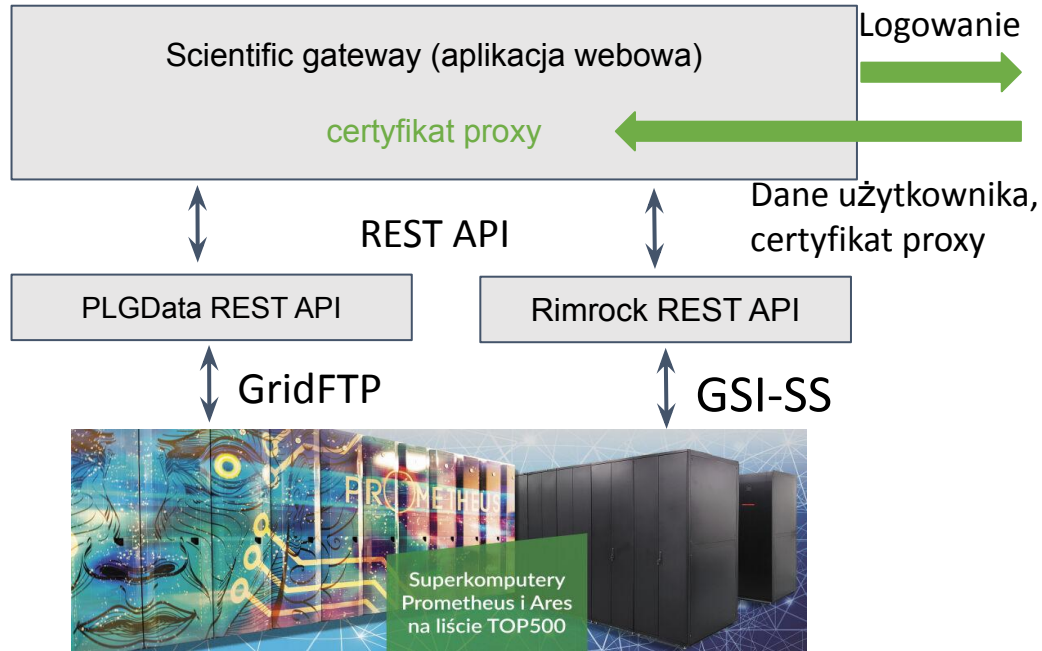
```
{
  "host": [string (required)],
  "working_directory": [string (optional)],
  "script": [string (required)]
}
```

. Below the JSON field, there is a warning message: "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." and an "Example:" section with a sample JSON object:

```
{
  "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}\nnecho hello\nexit 0"}' --header "Content-Type:application/json" --header "PROXY:$proxy" https://rimrock.plgrid.pl/api/jobs
```

Rozwiązanie!



Model Execution Environment

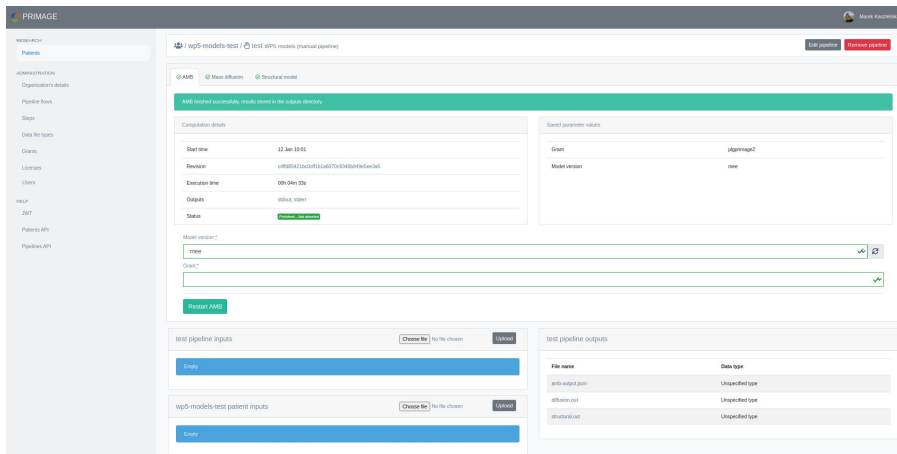
System dedykowany do wykonywania obliczeń, symulacji, podejmowania decyzji na temat pojedynczego przypadku oraz kohort.

Jak to działa?

- ❑ W momencie inicjowania danego przypadku tworzona jest dla niego dedykowana przestrzeń na HPC
- ❑ Każde obliczenie wykonywane jest w kontekście konkretnego przypadku
- ❑ Aby uruchomić obliczenie należy:
 - ❑ wybrać przypadek (lub kohortę)
 - ❑ kliknąć przycisk uruchamiający obliczenia
 - ❑ po zakończonych obliczeniach wyniki pojawiają się w dedykowanym katalogu, który nie jest nadpisywany przez inne obliczenia

Główne zalety:

- ❑ System zintegrowany z HPC dostępnymi w Cyfronecie (Prometeusz, Ares, Atena)
- ❑ Zarządzanie przypadkami, plikami wejściowymi, symulacjami oraz rezultatami w jednym miejscu
- ❑ Wersjonowanie symulacji dzięki integracji z GIT
- ❑ Automatyczne oraz manualne uruchomienie obliczeń



Model Execution Environment - uruchomienie obliczeń

1. System sprawdza, czy wszystkie wymagane dane wejściowe są dostępne, jeśli nie obliczenie nie może być uruchomione
2. Model zostaje pobrany z repozytorium GIT w wersji wybranej przez użytkownika
3. Na podstawie szablonu modelu tworzony jest model dedykowany dla konkretnego przypadku
4. System zleca zadanie na infrastrukturę HPC przy użyciu systemu kolejkowego slurm oraz serwisu Rimrock
5. Wymagane pliki wejściowe transferowane są na infrastrukturę HPC przy użyciu serwisu PLGData
6. MEE monitoruje uruchomione zadanie
7. w momencie zakończenia obliczenia rezultaty pobierane są z HPC i są bezpośrednio dostępne dla użytkownika

Rimrock

Robust Remote Process Control

Rimrock application simplify the way how you can interact with the to execute application as processes on a CLI node or in batch mode. a dedicated REST interface you will be able to start new job on

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.



```
2 #SBATCH --N 1
3 #SBATCH --output=/dev/null
4 #SBATCH --time=00:01:00
5 #SBATCH -A (( grant_id ))
6 #SBATCH -p plgrid-testing
7 #SBATCH --output /net/archive/groups/plgridimage/slurm_outputs/slurm-%j.out
8 #SBATCH --error /net/archive/groups/plgridimage/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 # Clone repository
19 (% clone_repo %)
20 echo -----END-----
21
22 echo Downloading numbers
23 (% stage_in_demo_numbers %)
24 echo -----END-----
25
26 echo Sorting
27 echo -----START-----
28 python demo-steps/1_sort.py
29 echo -----END-----
30
31 echo Uploading results
32 (% stage_out_steps.txt %)
33 echo -----END-----
34
35 exit 1
```

```
7 #SBATCH --N 1
8 #SBATCH --output=/dev/null
9 #SBATCH --time=00:01:00
10 #SBATCH -A plgridimg2
11 #SBATCH -p plgrid-testing
12 #SBATCH --output /net/archive/groups/plgridimage/slurm_outputs/slurm-%j.out
13 #SBATCH --error /net/archive/groups/plgridimage/slurm_outputs/slurm-%j.err
14
15 # Finish with error on first command with error
16 set -e
17
18 ## Change to the directory where sbatch was called
19 cd $SCRATCHDIR
20
21 ## Clone repository and switch into selected revision
22 echo Preparing computation source code
23 # Clone repository
24 (% clone_repo %)
25 echo -----END-----
26
27 echo Downloading numbers
28 (% stage_in_demo_numbers %)
29 echo -----END-----
30
31 echo Sorting
32 echo -----START-----
33 python demo-steps/1_sort.py
34 echo -----END-----
35
36 echo Uploading results
37 (% stage_out_steps.txt %)
38 echo -----END-----
39
40 exit 1
```


Wnioski

- ❑ HPC jest trudny, szczególnie dla osób domenowych (np. zajmujących się spersonalizowaną medycyną)
- ❑ Możemy pomóc dostarczając dedykowanych serwisów domenowych zintegrowanych z infrastrukturą HPC
- ❑ Przykłady pokazane na tej prezentacji:
 - ❑ Model Execution Environment - <https://mee.cyfronet.pl>
 - ❑ Rimrock - <https://rimrock.plgrid.pl>
 - ❑ PLGData - <https://data.plgrid.pl>

Chętnie weźmiemy udział w ambitnych projektach naukowych, zwłaszcza w ramach Horizon Europe

<https://dice.cyfronet.pl>
<https://dice-cyfronet.github.io>

