



IT Solutions and Hardware Resources extending the National Grid Infrastructure for Polish Science

J. Kitowski, K. Wiatr, Ł. Dutka,
T. Szepieniec, R. Pająk, M. Sterzel

ACK Cyfronet AGH

KU KDM 2016 17.03.2016

- PLGrid Program
- Achievements
- Selected services
- Summary

- Increasing importance of Computational Science and Big Data Analysis
- e-Science and e-Infrastructure interaction

■ Needs:

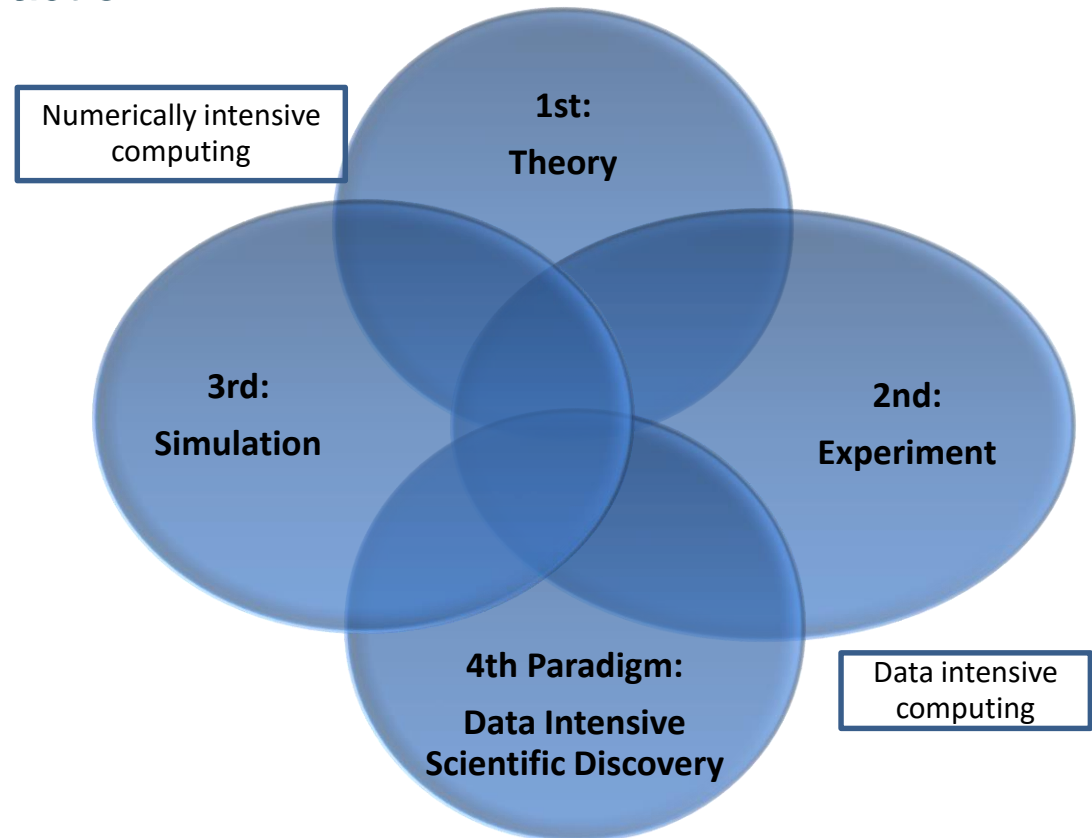
- Preventing users from technical problems
- Expert support for making science
- Increase of resources
- Openness for future paradigms

■ Experiments in silico:

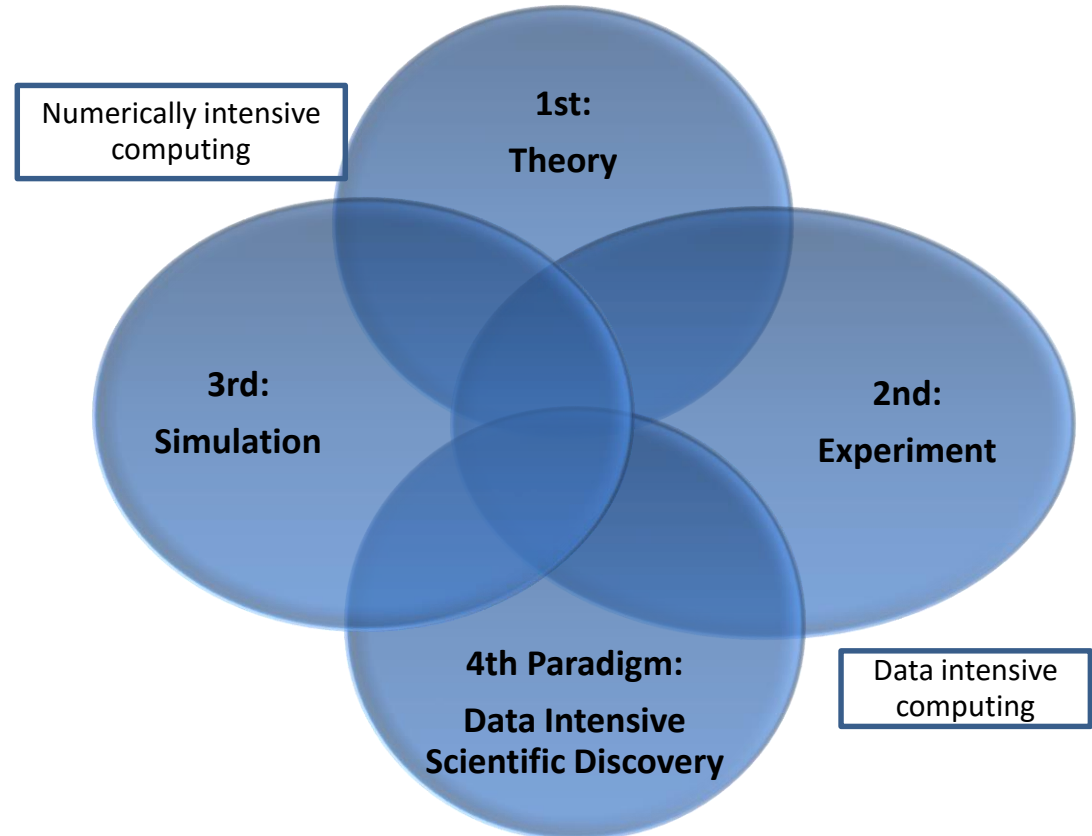
- Computing and Data for Open Science
- International collaboration
- User/platform driven e-infrastructure innovation

■ Computational Science problems:

- Algorithms, environments and deployment
- Future and emerging technologies

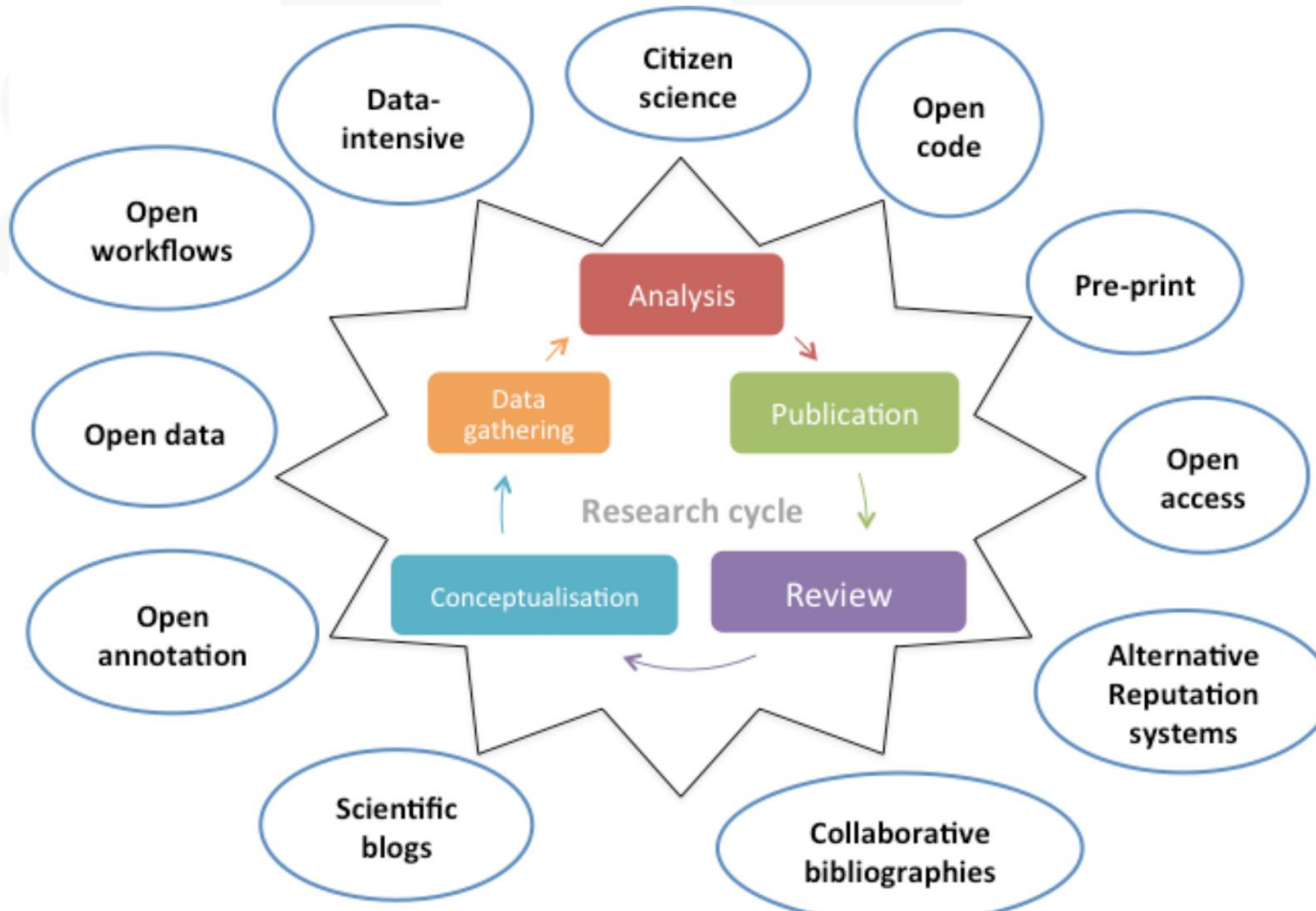


- Increasing importance of Computational Science and Big Data Analysis
- e-Science and e-Infrastructure interaction
-
-
-
- Open Science Commons



Open Science

Opening of the **creation** and **dissemination** of **scholarly knowledge** towards a multitude of stakeholders, from professional researchers to citizens



Openness
Participation
Collaboration
Sharing
Re-use



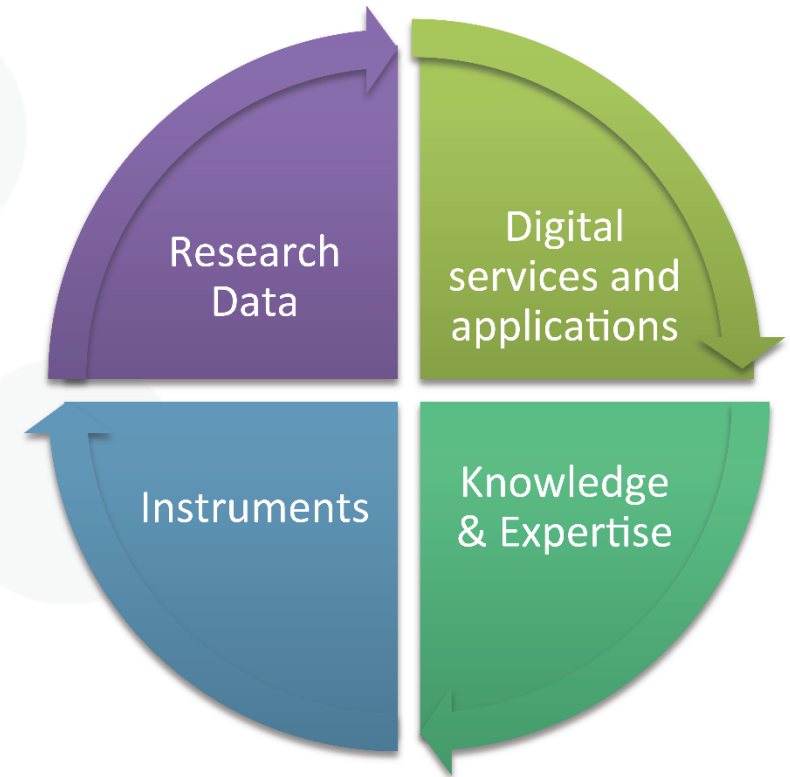
Greater social value

Source: <http://goo.gl/uO9MK5>

ACK: Sergio Andreozzi, EGI, EGI Conf. Lisbon, 2015

Open Science: a Complex Resource System

- Shared resources
 - Integrated, easy and fair access
- Engaged communities
 - Participating in the process
 - Culture of sharing
 - Collaborating in the management and stewardship
- Governance
 - Rules to access
 - Rules to resolve conflicts
 - Rules to balance quality vs. openness
- Financial support
 - For long-term availability



ACK: Sergio Andreozzi, EGI, EGI Conf. Lisbon, 2015



Akademickie Centrum Komputerowe
CYFRONET AGH w Krakowie
(koordynator)



Centrum Informatyczne Trójmiejskiej
Akademickiej Sieci Komputerowej



Poznańskie Centrum
Superkomputerowo
Sieciowe w Poznaniu

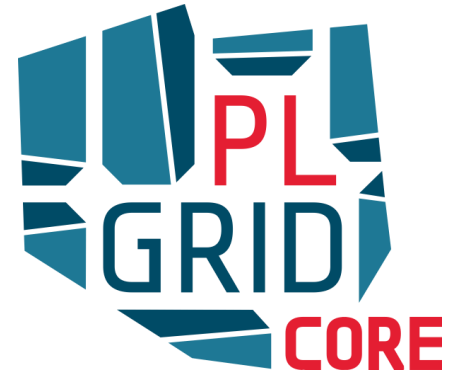


Interdyscyplinarne Centrum
Modelowania
Matematycznego i Komputerowego
w Warszawie



Wrocławskie Centrum
Sieciowo - Superkomputerowe we
Wrocławiu







ZASOBY OBLICZENIOWE

Rozproszone zasoby pięciu najważniejszych ośrodków superkomputerowych skoncentrowane w Infrastrukturze PLGrid składają się na najpotężniejszą infrastrukturę obliczeniową w Polsce.



PRZECHOWYWANIE DANYCH

Prowadzenie badań często wymaga analizy ogromnych ilości danych, które można bezpiecznie przechować i udostępnić PLGrid na przestrzeni pamięciowej ponad 60 PB.



PROFESJONALNE OPROGRAMOWANIE

Narzędzia i programy, które pozwolą na efektywne wykorzystanie zasobów Infrastruktury PLGrid oraz usprawnią prace badawcze wszystkich dziedzinach naukowych.



NARZĘDZIA PRACY ZESPOŁOWEJ

Implementacja biznesowych rozwiązań w nauce umożliwia realizację złożonych i długotrwałych projektów oraz ułatwia koordynację prac zespołów badawczych.



CHMURA OBLICZENIOWA

Innowacyjna odpowiedź na dynamiczny rozwój badań naukowych i zmianę specyfiki prac prowadzonych z użyciem zasobów obliczeniowych. Umożliwia optymalne dostosowanie środowiska obliczeniowego do prowadzonych badań.

Prometheus (ACK Cyfronet AGH)

- **39 miejsce na TOP500**
- **2,4 PFLOPS**, 53 568 rdzeni obliczeniowych, 279 TB pamięci RAM, 10 PB przestrzeni dyskowych

Bem (WCSS)

- **128 miejsce na TOP500**
- **0,86 PFLOPS**, 22 656 rdzeni obliczeniowych, 76,5 TB pamięci RAM, 1,1 PB przestrzeni dyskowych

Zeus (ACK Cyfronet AGH)

- **387 miejsce na TOP500**
- **0,37 PFLOPS**, 25 468 rdzeni obliczeniowych, 52 TB pamięci RAM, 800 TB przestrzeni dyskowych

Topola (ICM UW)

- **0,24 PFLOPS**, 6244 rdzeni obliczeniowych, 18 TB pamięci RAM, 430 TB przestrzeni dyskowych

Eagle (PCSS)

- **80 miejsce na TOP500**
- **1,37 PFLOPS**, 32 984 rdzeni obliczeniowych, 120 TB pamięci RAM, 47 PB przestrzeni dyskowych

Tryton (CI TASK)

- **163 miejsce na TOP500**
- **0,64 PFLOPS**, 38 568 rdzeni obliczeniowych, 218 TB pamięci RAM, 256 TB przestrzeni dyskowych

Inula (PCSS)

- **0,30 PFLOPS**, 6 944 rdzeni obliczeniowych, 19 TB pamięci RAM, 43 TB przestrzeni dyskowych

Hydra (ICM UW)

- **0,10 PFLOPS**, 5 056 rdzeni obliczeniowych, 22 TB pamięci RAM, 430 TB przestrzeni dyskowych



- 5+ PFLOPS
- 90 000+ rdzeni
- 700+ TB RAM
- 40+ PB dysków



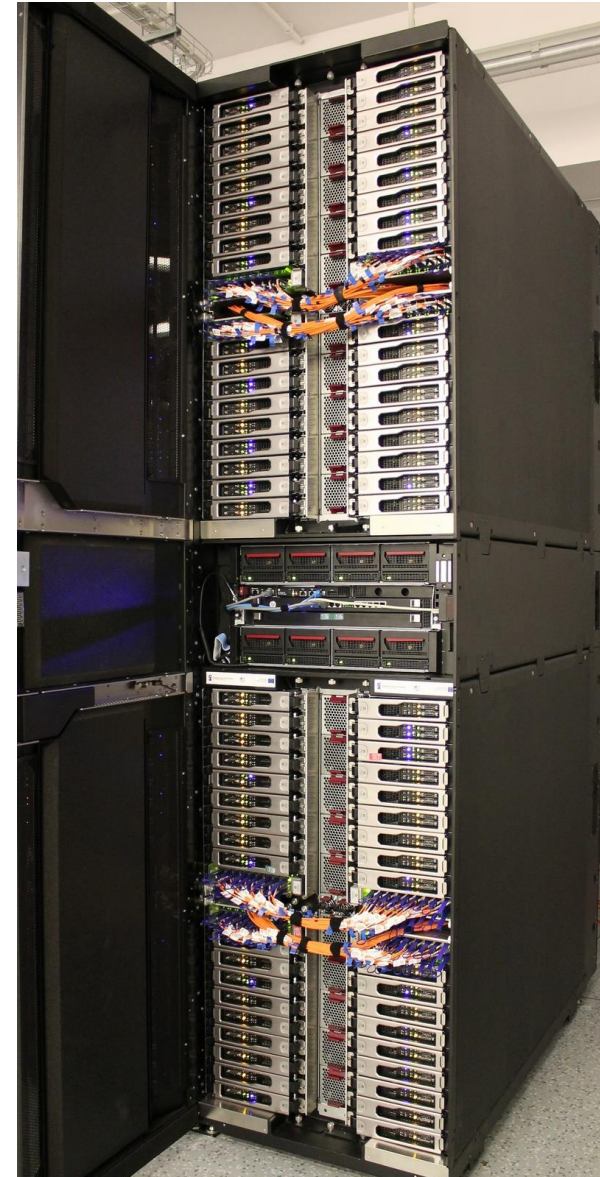
■ Lista systemów z Polski

- **49** - Prometheus (ACK Cyfronet AGH) - 1,65 PFLOPS (PLGrid)
- **126** – Tryton (TASK) – 0,63 PFLOPS (część dostępna w PLGrid)
- **135** – Bem (WCSS) – 0,63 PFLOPS (część dostępna w PLGrid)
- **155** – Centrum Informatyczne Świerk – 0,49 PFLOPS
- **269** – Zeus (ACK Cyfronet AGH) – 0,37 PFLOPS (PLGRrid)
- **380** – Orion (ICM) – 0,20 PFLOPS
- **418** – Nostromo (ICM) – 0,19 PFLOPS

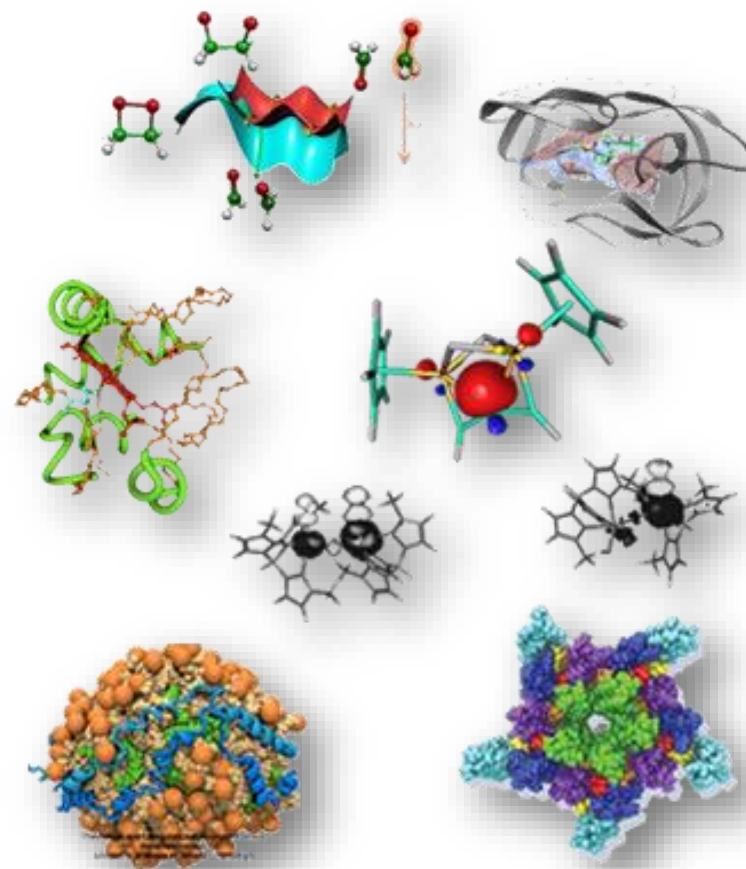
■ Lista systemów z Polski

- **39** – Prometheus (ACK Cyfronet AGH) – 2,35 PFLOPS (PLGrid)
- **80** – Orzeł/Eagle (PCSS) – 1,37 PFLOPS (PLGrid)
- **128** – Bem (WCSS) – 0,86 PFLOPS (część dostępna w PLGrid)
- **163** – Tryton (TASK) – 0,63 PFLOPS (część dostępna w PLGrid)
- **218** – Centrum Informatyczne Świerk – 0,49 PFLOPS
- **387** – Zeus (ACK Cyfronet AGH) – 0,37 PFLOPS (PLGRrid)

- HP Apollo 8000
- 2,35 PFLOPS (**39** na Top500 XI 2015)
- 2232 węzłów, 2x Intel Haswell E5-2680v3 2,5 GHz, 12 rdzeni
 - w tym 72 węzły z 2 x Tesla K40 XL
- 53 568 rdzeni, 13824 na wyspie
- 279 TB DDR4 RAM (128 GB na węzeł)
- PUE ~1.05, 816 kW mocy całkowitej
- 20 szaf typu rack
- CentOS 7
- system gotowy do rozbudowy w trakcie produkcyjnego działania



- **Chemia i Biologia:** ADF, AMBER, CFOUR, Dalton, GAMESS, Gaussian, Molcas, Molpro, MOPAC, NWChem, Open Babel, TURBOMOLE, AutoDock/AutoGrid, BLAST, Clustal, Siesta, Quantum Espresso, VASP (na licencjach własnych użytkowników), CPMD, Gromacs, NAMD
- **FEM:** Abaqus, Ansys/FLUENT, Nastran, OpenFOAM
- **Nanotechnologia:** ABINIT, Quantum Espresso, NAMD
- **Interdyscyplinarne:** Mathematica, MATLAB, Simulink
- **Kompilatory i narzędzia:** Intel, PGI, GNU compilers, MKL, CUDA, MPI, OpenMP, Alinea, Python, R, Ruby
- **Bazy danych**
- **Możliwość instalacji dowolnego oprogramowania**



<https://aplikacje.plgrid.pl/>

Adobe Connect - to cenione na rynku rozwiązanie do prowadzenia telekonferencji do 50 uczestników jednocześnie.

telekonferencje

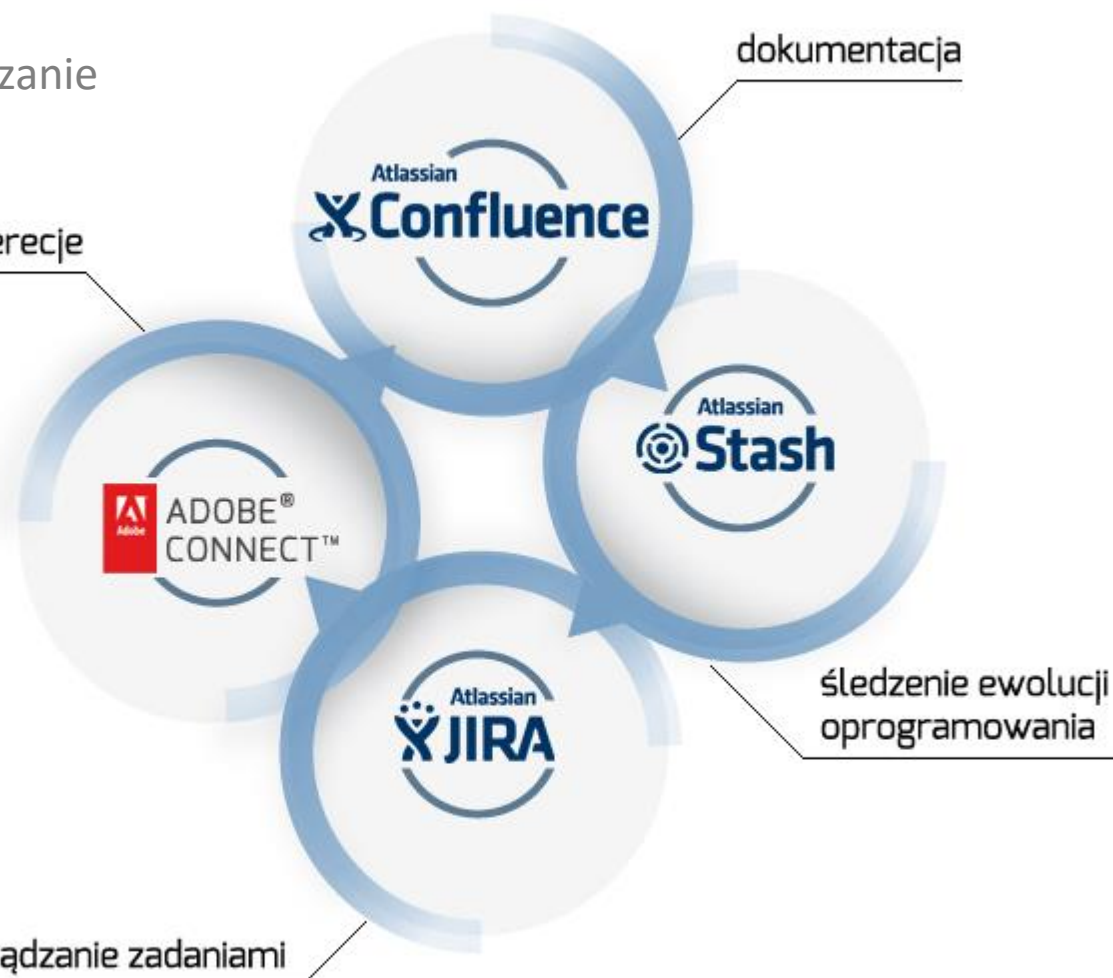
Confluence - to przestrzeń typu "wiki" do agregowania wiedzy.

Jira - to narzędzie do planowania pracy i śledzenia jej efektów.

Stash - oprogramowanie do śledzenia zmian w kodzie źródłowym


zarządzanie zadaniami















dokumentacja



<https://data.plgrid.pl/>

- Dostęp do danych przez przeglądarkę internetową

ZAWARTOŚĆ FOLDERU: PEOPLE /  / BADANIA / LUTY 2014 / [Dodaj pliki](#) [Nowy folder](#)

Prawa	Rozmiar	Data modyfikacji	Nazwa	Typ	
-rw-r--r-- 	11.6 MB	Jan 27 22:24	 GSM1019938_NBM6-P.CEL	Plik	Na pewno? Tak Nie
<input type="text" value="rw-r-----"/>  	3.5 MB	Jan 27 22:31	 GSM933008_US91903683 Dec08_1_1.txt	Plik	Usun
-rw-r--r-- 	5.11 MB	Feb 5 16:25	<input type="text" value="GSM980568_MB2009031253.CEL"/>  	 Plik 	Usun
drwxr-xr-x 	2 KB	Feb 8 14:47	 aktualne_pliki	Folder	Usun
-r--r-- 	128 KB	25 21:07	 all_new-3class.tr	Plik	Usun

<https://submit.plgrid.pl/>



- Rimrock – Robust Remote Process and Job Controller
- Ułatwia pracę z zadaniami obliczeniowymi na zdalnych zasobach
- Umożliwia pracę z zadaniami wsadowymi oraz uruchomienie interaktywnej aplikacji
 - dostarcza prostego interfejsu RESTowego do wysyłania i odbioru danych oraz kontroli zadań

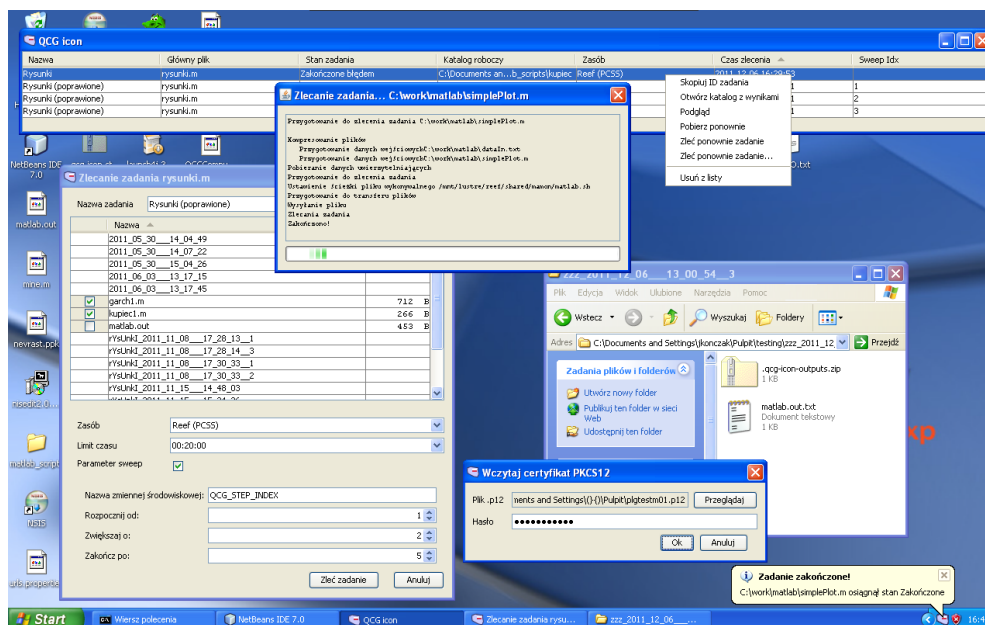


<http://www.qoscosgrid.org/trac/qcg-icon>

- Lekka aplikacja graficzna do zarządzania zadaniami gridowymi
- Szerokie portfolio wspieranych aplikacji z chemii obliczeniowej

- Gaussian
- GAMESS
- Molpro
- Dalton
- Turbomole
- ADF
- CRYSTAL09

- integracja z GaussView





<http://insilicolab.grid.cyfronet.pl/>

- Środowisko pracy z systemem zintegrowanych narzędzi, które:
 - wspomagają zarządzanie złożonymi obliczeniami
 - automatyzują powtarzalne cykle obliczeń
 - umożliwiają w wygodny sposób zarządzanie procesem obliczeń
 - ułatwiają zarządzanie rozproszonymi danymi eksperymentu
 - umożliwiają wspólną analizę rezultatów wielu równoległych obliczeń
 - ułatwiają współpracę pomiędzy badaczami pracującymi nad wspólnymi projektami
 - nie rozpraszają użytkowników wykorzystywaną technologią - bez forsowania zmiany sposobu myślenia naukowców



Your Experiments

Menu



✓ Benzene...ground
state...RHF/STO-2G

Log in... You are logged in as **anonymousUser**

Welcome



Welcome to InSilicoLab Portal

To take advantage of the full portal functionality (including job submission and LFC directory management), it is required that you have a valid proxy configured. You can configure it now - using the left panel, or do it later at any time - by clicking the "Configure proxy..." button in the workspace.

If you choose not to configure your proxy now, you can still access the portal and work with limited functionality. For this purpose, choose one of the options right panel below.

Actions

- Configure your proxy**
- Create a new experiment:**

<http://insilicolab.grid.cyfronet.pl/>



■ Chemia kwantowa oraz biochemia

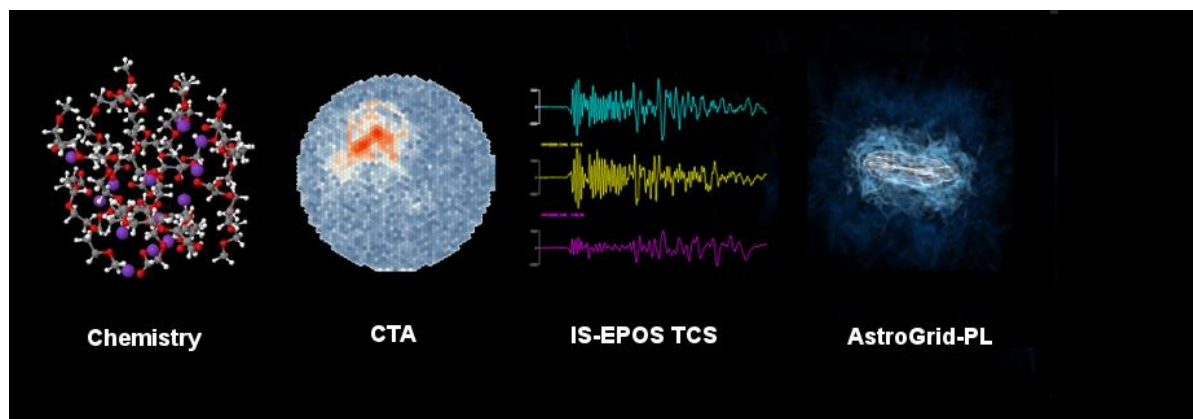
- obliczenia pakietami Gaussian, GAMESS, Turbomole
- Trajectory Sculptor - półautomatyczne przycinanie dużych układów molekularnych (np. wyników symulacji MD)
- możliwość łączenia obu typów eksperymentów

■ Astrofizyka

- obliczenia hydrodynamiczne metodami objętości skończonej
- obliczenia dla konsorcjum Cherenkov Telescope Array (CTA)

■ Geofizyka

- badania sejsmiczności indukowanej (IS-EPOS)



<https://statystyka.plgrid.pl/>

- Zintegrowana platforma usług dedykowanych statystycznej analizie danych w wielodziedzinowych zagadnieniach naukowych
 - ułatwia wybór testu statystycznego dla analizy badanego zagadnienia
 - udostępnia otwarte repozytorium kilkudziesięciu popularnych testów statystycznych
 - oraz zestaw aplikacji do analizy szeregów czasowych.

Statystyczna Analiza Danych ▾ Kategorie ▾ Dodaj test Dashboard Tomasz Piontek ▾

[Lista testów](#) / [Jednoczynnikowa analiza wariancji \(ANOVA\)](#) / [domyślna](#) / [Uruchomienie testu](#)

Jednoczynnikowa analiza wariancji (ANOVA)

Wyniki **Dane** Raport

wpisz lub wklej z Excela dane, oddzielając liczby spacjami.

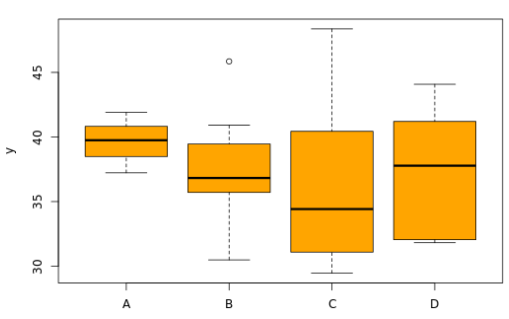
wpisz lub wklej z Excela podział na grupy

załaduj plik Excela

No file selected.

Opis osi y

Zmienna 1: kolor:



Porównanie grup przy pomocy jednoczynnikowej analizy wariancji: $p=0.831$ wartość statystyki $F=0.291$ na 3 i 36 stopniach swobody

Wyjaśnienia

Program do porównywania grup przy pomocy analizy wariancji (ANOVA). Należy podać kolumnę danych z wartościami numerycznymi, oraz, w drugim polu, kolumnę z podziałem na grupy (do której grupy dana wartość należy).

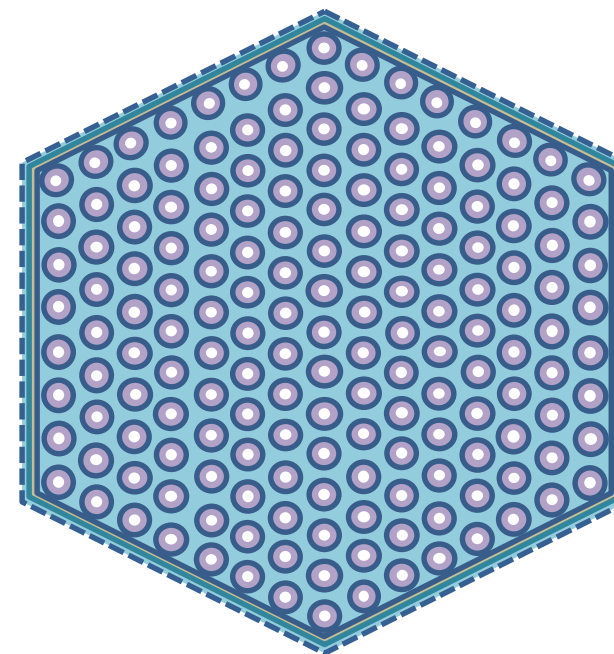
Dane można wprowadzać na trzy sposoby:

- wpisując dane bezpośrednio do okienek po lewej stronie aplikacji
- w Excelu lub OpenOffice podświetlamy kolumnę (kolumny nie powinny mieć nazw), Ctrl+c i po kliknięciu okienka po lewej stronie, Ctrl+v
- wybierając przy pomocy przycisku po lewej stronie plik Excela, który w pierwszym arkuszu, pierwszej kolumnie ma wartości liczbowe a w drugiej podział na grupy. Kolumny nie powinny mieć nazw.



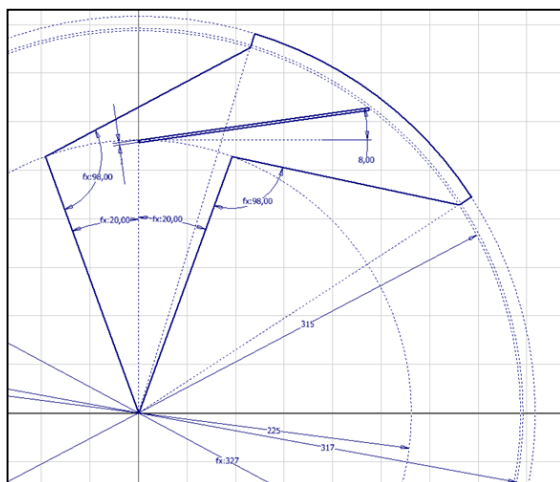
The Monte Carlo Continuous
Energy Burn-up Code

- Symulacje wywołanych promieniowaniem jądrowym zmian w materii
- Przykłady użycia
 - Analiza bezpieczeństwa reaktorów jądrowych
 - Reaktory jądrowe IV Generacji
 - Detektory promieniowania
 - Osłony radiacyjnych
 - Systemy podkrytyczne
 - Jądrowy cykl paliwowy
 - Elementy systemów składowania oraz przetwarzania zużytego paliwa jądrowego
 - Eksperymenty reaktorowe
 - Dozymetria

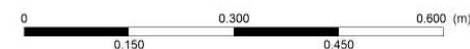
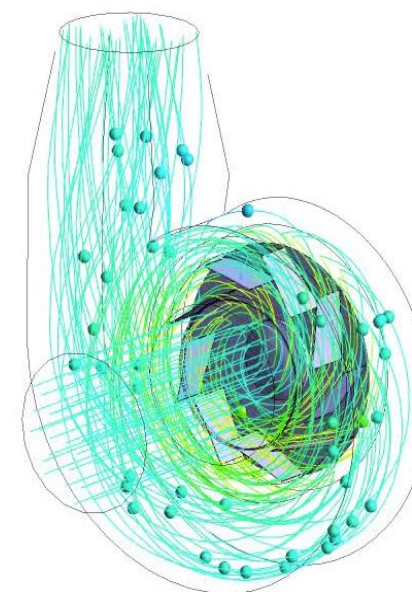


- Zintegrowany system do projektowania i optymalizacji maszyn rotodynamicznych

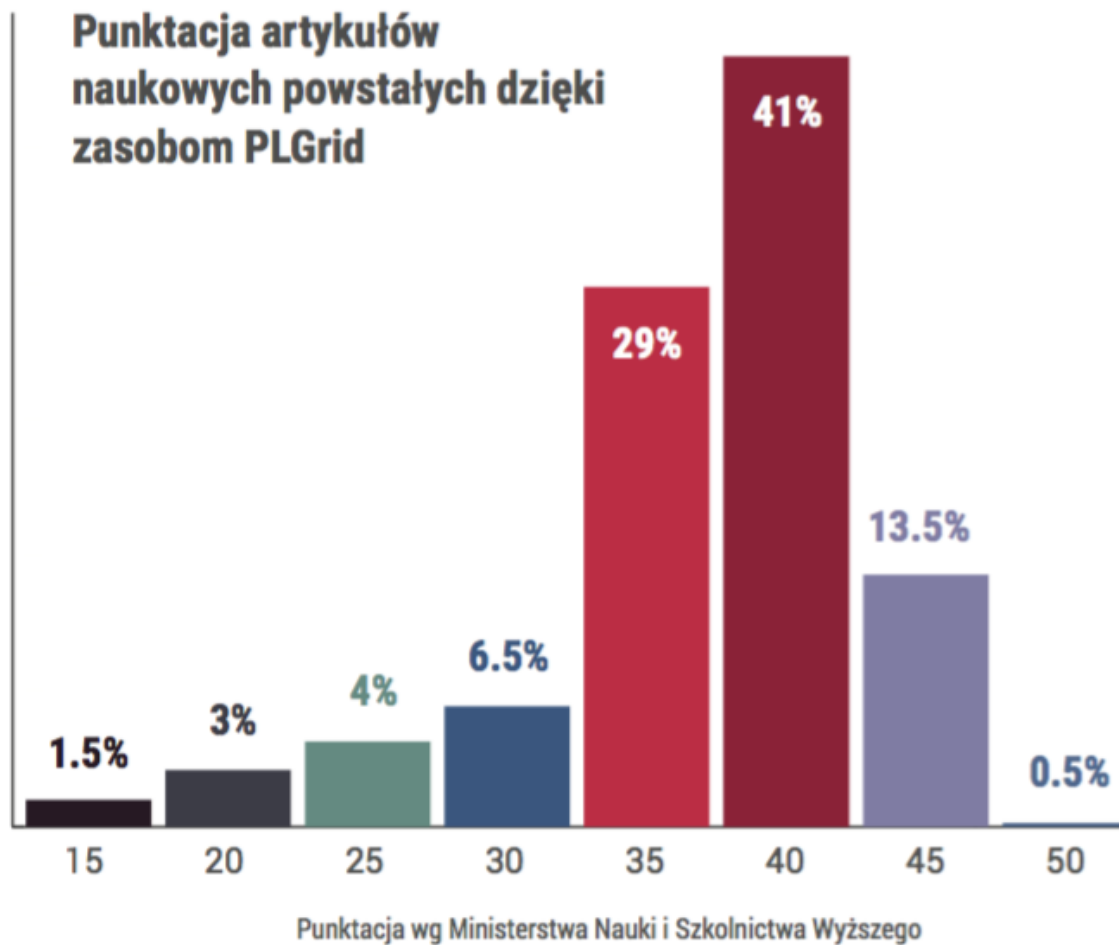
Od koncepcji



do maszyny



Punktacja artykułów naukowych powstałych dzięki zasobom PLGrid



Dostęp do Infrastruktury



- Rejestracja elektroniczna przez Portal Użytkownika PLGrid (<https://portal.plgrid.pl>)
- Weryfikacja użytkownika na podstawie nr OPI i danych w bazie Ludzie Nauki (<http://www.nauka-polska.pl/>)



Po rejestracji użytkownik uzyskuje możliwość:

- Zarządzanie dostępem do usług i wielu aplikacji
- Monitoring zużycia zasobów
- Zarządzanie grantami obliczeniowymi
- Zarządzanie certyfikatami gridowymi
- Dostęp do wszystkich zasobów przy użyciu **jednego konta i jednego hasła** (lub certyfikatu gridowego)

REJESTRACJA

1

Wypełnij gotowy formularz rejestracji i wyślij go do weryfikacji. Następnie kliknij w link aktywacyjny, który otrzymałeś w mailu. Konto zostanie aktywowane!

AFILIACJE

2

Uzupełnij informacje na temat instytucji naukowej, do której przynależysz. Afiliacja polskiej jednostki naukowej umożliwi dostęp do pełnej oferty Infrastruktury PLGrid.

ZAŁÓŻ ZESPÓŁ NAUKOWY

3

Jest to jednostka, za pomocą której możliwe jest negocjowanie zasobów obliczeniowych. Odzwierciedla cały zespół badawczy lub użytkownika pracującego samodzielnie.

UTWÓRZ GRANT

4

Grant stanowi umowę między centrum komputerowym a użytkownikiem. Określany jest w niej rodzaj usługi, jej parametry oraz gwarancje otrzymania.

- PLGrid to różne sposoby dostępu do zasobów
 - lokalne systemy kolejkowe na klastrach (CLI oraz GUI)
 - oprogramowania pośredniczące (UNICORE, QosCosGrid, gLite)
 - maszyny wirtualne
 - portale sieciowe (InSilicoLab, GridSpace2, PLG-Data, DataNet,...)
- Zarządzanie dostępem do usług poprzez Portal Użytkownika PLGrid (<https://portal.plgrid.pl>)
- Katalog usług
 - usługi ogólne - <https://docs.plgrid.pl/uslugi>
 - usługi dedykowane wybranym dziedzinom nauki - https://docs.plgrid.pl/uslugi_dziedzinowe

Szkolenia

- zdalne - <https://portal.plgrid.pl/> lub <https://ekursy.plgrid.pl/>
- z trenerem - informacje na <http://www.cyfronet.krakow.pl/>

Podręcznik Użytkownika

- Dostępny online: https://docs.plgrid.pl/podrecznik_uzytkownika

Konsultacje z ekspertami

- Za pośrednictwem Helpdesk PLGrid: <https://helpdesk.plgrid.pl>
helpdesk@plgrid.pl

Forum użytkowników

- <https://zapytaj.plgrid.pl/>

- Three dimensions of development:
 - HPC/GRID/CLOUDs
 - Data & Knowledge layer
 - Network & Future Internet
- International collaboration
- Deployments have the national scope; however with close European links
- Development oriented on end-users & research projects
- Achieving synergy between research projects and e-infrastructures by close cooperation and offering relevant services
- Durability at least 5 years after finishing the projects - confirmed in contracts
- Future plans: continuation of development
- Center of Excellence
- CGW, KUKDM as places to exchange experience and for collaboration between eScience centers in Europe

■ ACC Cyfronet AGH

- Kazimierz Wiatr
- Michał Turała
- Marian Bubak
- Krzysztof Zieliński
- Łukasz Dutka
- Tomasz Szepieniec
- Marcin Radecki
- Mariusz Sterzel
- Karol Krawentek
- Agnieszka Szymańska
- Maciej Twardy
- Angelika Zaleska-Walterbach
- Andrzej Oziębło
- Renata Słota
- Tomasz Gubała
- Darin Nikolow
- Aleksandra Pałuk
- Patryk Lasoń
- Marek Magryś
- Łukasz Flis
- Robert Pająk
- Klemens Noga

Special thanks to
many
domain experts !

... and many others.....

■ ICM

- Marek Niezgódka
- Piotr Bała
- Maciej Filocha

■ PCSS

- Maciej Stroiński
- Norbert Meyer
- Krzysztof Kurowski
- Tomasz Piontek
- Paweł Wolniewicz

■ WCSS

- Jacek Oko
- Józef Janyszek
- Mateusz Tykierko
- Paweł Dziekoński
- Bartłomiej Balcerek

■ TASK

- Rafał Tylman
- Mścisław Nakonieczny
- Jarosław Rybicki



Rejestracja: <https://portal.plgrid.pl>

helpdesk@plgrid.pl

+48 12 632 3355 wew. 306 lub bezpośrednio +48 12 632 5163