



Efektywne wykorzystanie zasobów PLGrid w chemii obliczeniowej

Klemens Noga

ACK Cyfronet AGH

Wydział Chemii Uniwersytetu Jagiellońskiego, Kraków 23 X 2015

- ACK Cyfronet AGH – główne kompetencje
- Infrastruktura PLGrid
 - PLGrid to ...
 - Struktura
 - Oferta
- Dostęp do Infrastruktury
 - Kto może uzyskać dostęp
 - Portal PLGrid
 - Wnioskowanie o zasoby
- Dlaczego my .-)

ACK Cyfronet AGH



Organizacja międzynarodowych i krajowych konferencji naukowych,

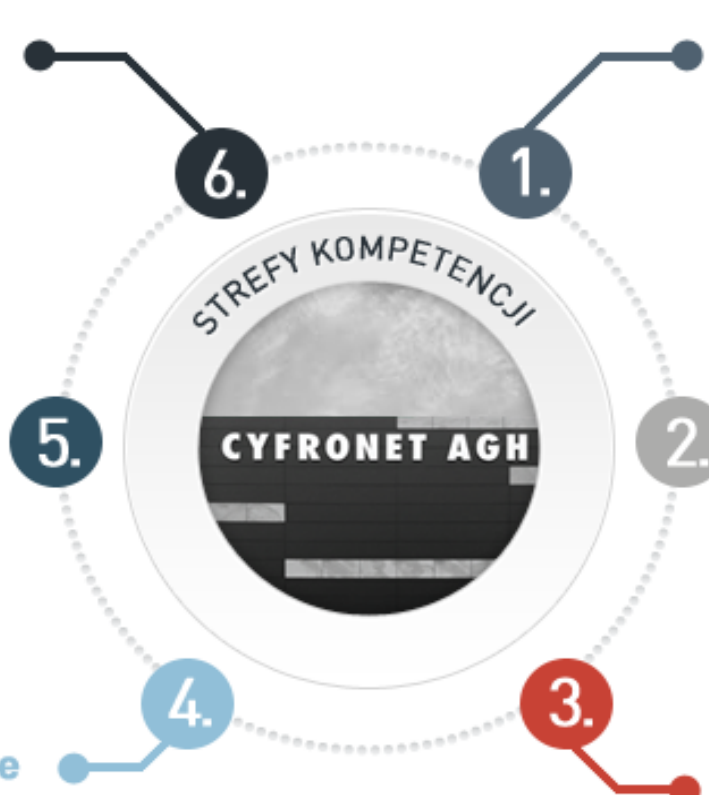
poświęconych tematyce informatycznej.

e-Nauczanie

efektywny system Blackboard (nauczania poprzez sieć) dla środowiska akademickiego.

Badania naukowe

prowadzone w ramach szeregu projektów badawczych, zarówno krajowych jak i międzynarodowych.



Miejska Sieć Komputerowa w Krakowie

rozwój i utrzymanie MANu w Krakowie oraz zapewnienie szerokiego spektrum usług sieciowych dla użytkowników.

Komputery dużej mocy

udostępnianie środowisku naukowemu mocy obliczeniowych, zasobów pamięci dyskowych i taśmowych.

Oprogramowanie specjalistyczne

dostęp do oprogramowania naukowego i specjalistycznego dla użytkowników komputerów dużej mocy.

Infrastruktura **PL**Grid



**Infrastruktura obliczeniowa
dla polskich naukowców.
Udostępnia usługi informatyczne
o satysfakcjonującym
i gwarantowanym
poziomie jakości**



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



W infrastrukturze liczy blisko **3000 użytkowników**.

Efektom prac prowadzonych z pomocą Infrastruktury jest ponad **1500 publikacji**.

Znakomita większość publikacji jest znajduje się w czasopismach z **Listy Filadelfijskiej** oraz jest wysoko punktowana przez **MNiSW**

Z infrastruktury korzystają:

- Pracownicy naukow i dydaktyczni
- Doktoranci
- Magistralni w ramach prac dyplomowych
- Studenci
- Inne osoby uczestniczące w badaniach naukowych (w tym obcokrajowcy)

Warto wiedzieć o PLGrid

1000

nowych użytkowników
w ciągu 6 miesięcy

1,7

PetaFlopsa
więcej mocy
od 28.04.2015

1515

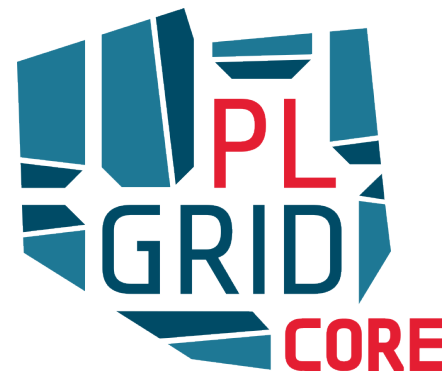
punktowanych publikacji naukowych
powstałych przy wykorzystaniu
Infrastruktury PLGrid

WAŻNE WYDARZENIA





Dziedzinowe Usługi Nowej
Generacji w Infrastrukturze
PLGrid dla Polskiej Nauki



Centrum Kompetencji
w Zakresie Rozproszonych
Infrastruktur obliczeniowych
Typu Gridowego - PLGridCore

Oferta



Zasoby obliczeniowe



Przechowywanie
danych



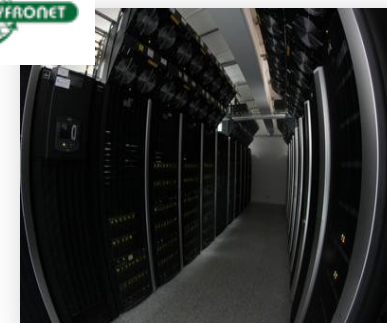
Profesjonalne
oprogramowanie



Narzędzia pracy
zespołowej



Chmura obliczeniowa



- 588 TFLOPS
- 40 288 rdzeni
- 108 TB RAM
- 5.8 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

- HP Apollo 8000
- 1.65 PFLOPS (**49** na Top500 VI 2015)
- 1728 węzłów, 2x Intel Haswell E5-2680v3 2,5 GHz, 12 rdzeni
- 41 472 rdzeni, 13 824 na wyspie
- 216 TB DDR4 RAM (128 GB na węzeł)
- PUE ~1.05, 680 kW mocy całkowitej
- 15 szaf typu rack, 13 m²
- CentOS 7
- system gotowy do rozbudowy w trakcie produkcyjnego działania
- rozbudowa o czwartą wyspę obliczeniową (IX-X 2015)
 - dodatkowe 504 węzły (i ponad 0,5 PFLOPS)
 - w tym 72 węzły z 2 x Tesla K40





- HP Apollo 8000
 - efektywne chłodzenie wodne
 - wydajność energetyczna (~1,05 PUE)



- 2,5+ PFLOPS
- 85 000+ rdzeni
- 300+ TB RAM
- 15+ PB dysków



Dostępne rodzaje procesorów

Intel Xeon 4-, 6-, 12-, 14-rdzeniowe (do 28 rdzeni na węźle)

AMD Opteron 6-, 12-, 16-rdzeniowe (do 64 rdzeni na węźle)

Akceleratory obliczeń

karty GPGPU NVidia Tesla (do 8 kart na węźle)

Intel Xeon Phi (do 2 kart na węźle)

Różne konfiguracje węzłów obliczeniowych

od 8 do 64 rdzeni na węzeł

do 512 GB RAM na węzeł

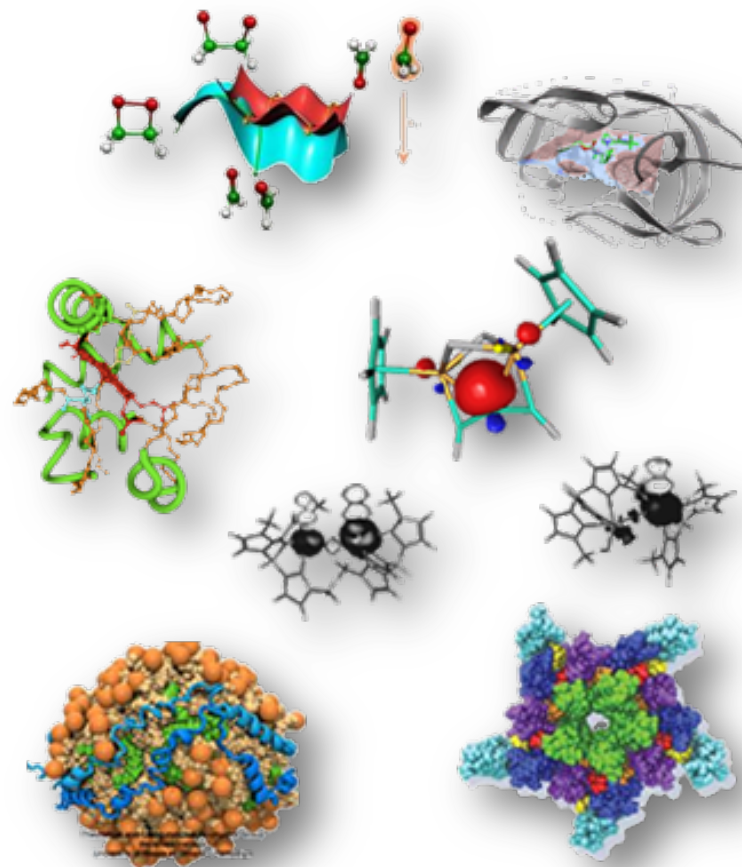
vSMP (Intel Xeon) – do 6 TB RAM oraz 768 rdzeni

Pełna konfiguracja na stronie:

www.plgrid.pl/oferta/zasoby_obliczeniowe/opis_zasobow/HPC

- PLGrid to różne sposoby dostępu do zasobów
 - lokalne systemy kolejkowe na klastrach (CLI oraz GUI)
 - oprogramowania pośredniczące (UNICORE, QosCosGrid, gLite)
 - maszyny wirtualne (<https://cloud.plgrid.pl>)
 - 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

- **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"/>	 	Usun
drwxr-xr-x 	2 KB	Feb 8 14:47	 aktualne_pliki	Folder	Usun
-r--r-- 	128 KB	25 21:02	 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

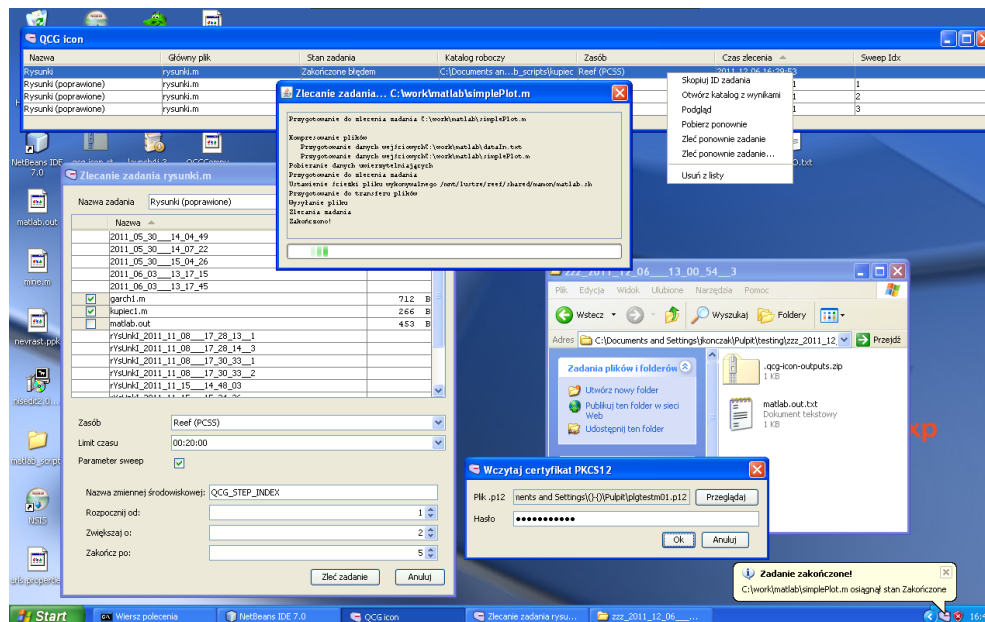
<http://www.qoscosgrid.org/trac/qcg-icon>



- 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://insicolab.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 rozprasza użytkowników wykorzystywaną technologią - bez forsowania zmiany sposobu myślenia naukowców




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

Your Experiments

Menu    

- ✓ Benzene...ground
state...RHF/STO-2G

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://insicolab.grid.cyfronet.pl>



■ Chemia kwantowa oraz biochemia

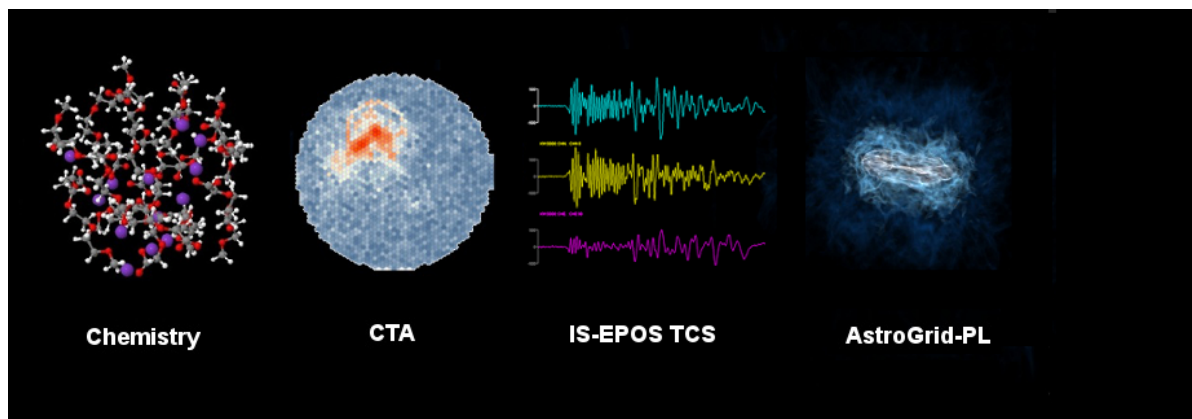
- obliczenia pakietami Gaussian, GAMESS, Turbomole
- Trajectory Sculptor - półautomatyczne przycinanie dużych układów molekularnych (np. wyników symulacji MD)
- możliwość łączenia obu typów eksperymentów

■ Astrofizyka

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

■ Geofizyka

- badania sejsmiczności indukowanej (IS-EPOS)



<https://insilicolab.chemia.plgrid.pl/>

- Dwa typy eksperymentów
 - obliczenia z wykorzystaniem metod chemii kwantowej
 - Trajectory Sculptor

- Obliczenia metodami chemii kwantowej

- dostępne aplikacje

- Gaussian

- GAMESS

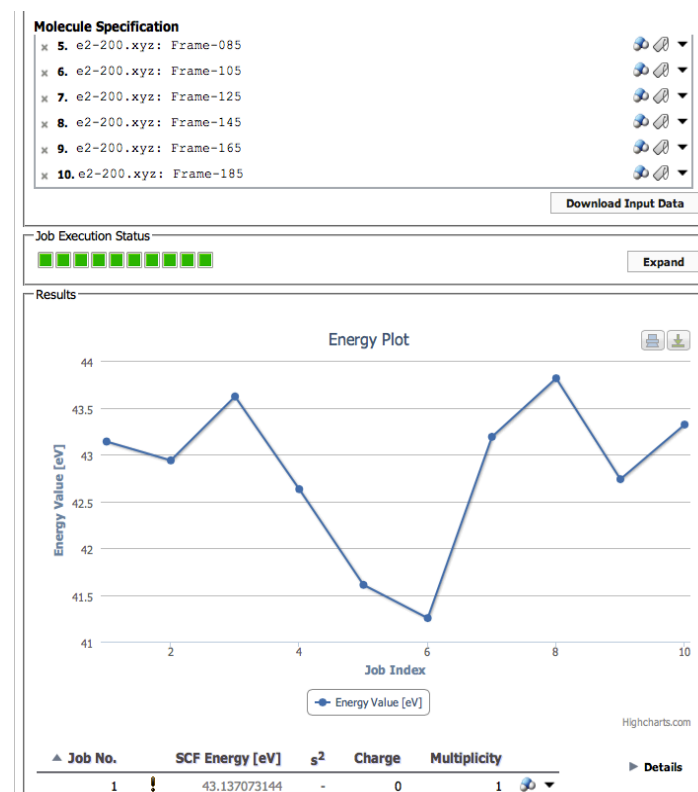
- TUROBMOLE

- Niedoida

- Terachem

- automatyczna wstępna analiza wyników obliczeń

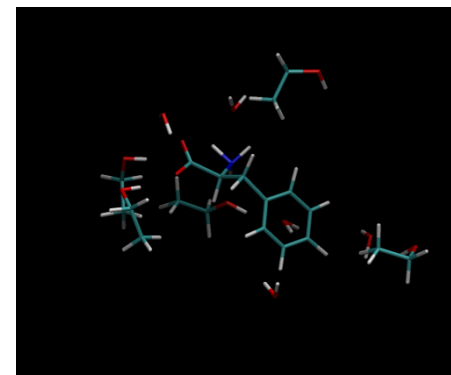
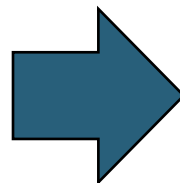
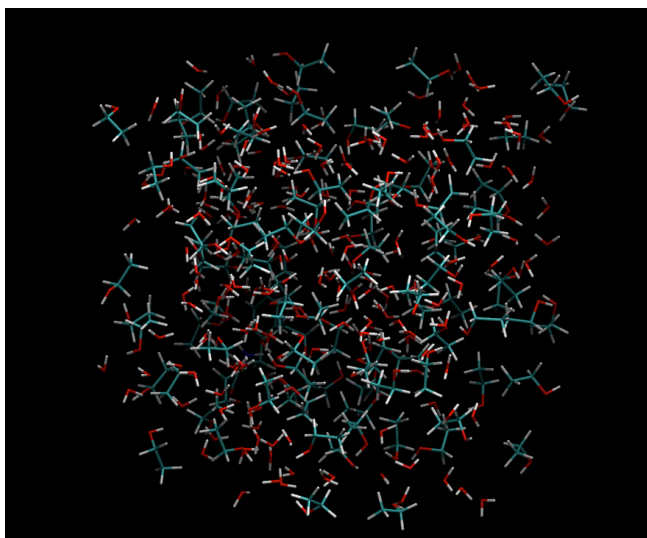
- możliwość równoczesnych obliczeń dla różnych geometrii w jednym eksperymencie





<https://insilicolab.chemia.plgrid.pl/>

- **Trajectory Sculptor** – narzędzie do przetwarzania trajektorii wynikowych z obliczeń dynamiki molekularnej
 - automatyczna ekstrakcja istotnych dla użytkownika fragmentów struktur z trajektorii MD
 - automatyczne przycinanie ramek trajektorii wybranych przez użytkownika
 - wyniki eksperymentu mogą być użyte w obliczeniach wykorzystujących metody chemii kwantowej



<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)

wpisz lub wklej z Excela dane, oddzielając liczby spacjami.

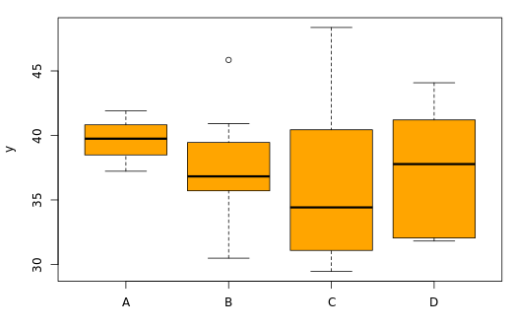
wpisz lub wklej z Excela podział na grupy

załaduj plik Excela
 No file selected.

Opis osi y

Zmienna 1: kolor:

Wyniki **Dane** Raport



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.

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)

- Certyfikat jest niezbędny do dostępu do infrastruktury gridowej, jest „dowodem osobistym” użytkownika i potwierdza jego tożsamość w usługach gridowych – zastępuje hasło
- Certyfikaty są wystawiane przez zaufane Centra Certyfikacji (ang. *Certification Authority, CA*)
- Certyfikaty zwykle ważne są **jeden rok**
- Zadania obliczeniowe posługują się certyfikatem krótko żyjącym (proxy)
 - możliwe jest jego automatyczne odnawianie



- Dla użytkowników PL-Gridu dostępne są dwa CA
 - Simple CA (<http://plgrid-sca.wcss.wroc.pl>)
 - Polish Grid CA (<https://plgrid-ca.pl>)
- Simple CA:
 - certyfikat uzyskuje się poprzez portal <https://portal.plgrid.pl/>
 - generowany na żądanie dla każdego użytkownika (automatycznie)
 - dostęp ograniczony do polskich zasobów
- PL-Grid CA
 - certyfikat uzyskuje się poprzez portal <https://plgrid-ca.pl>
 - użytkownik musi potwierdzić tożsamość w Urzędzie Rejestracji (RA)
 - umożliwia pracę na całości gridu w European Grid Initiative (EGI)

- Certyfikaty gridowe mogą być przechowywane w różnych formatach. Najbardziej rozpowszechnione to:
 - PKCS #12
 - certyfikat wraz z kluczem prywatnym znajduje się w jednym pliku binarnym zwykle o rozszerzeniu `.p12`
 - PEM
 - certyfikat stanowi para plików tekstowych:
 - klucz prywatny (zwykle `userkey.pem`)
 - plik certyfikatu (zwykle `usercert.pem`)
- Certyfikaty w formacie PEM używane są
 - większość oprogramowania pośredniczącego (często domyślny format)
- Certyfikaty w formacie PKCS #12 używane są przez
 - przeglądarki internetowe
 - większość oprogramowania pośredniczącego

Recepta:

- Wygeneruj swój certyfikat gridowy
 - np. certyfikat z Simple CA na Portalu Użytkownika PL-Grid
 - zakładka "Moje Konto", ramka "Certyfikaty Użytkownika"
- Zainstaluj swój certyfikat gridowy w przeglądarce
- Przejdź na Portal Użytkownika PL-Grid - <https://portal.plgrid.pl>
- Zarejestruj swój certyfikat w Portalu PL-Grid
 - zakładka "Moje Konto", ramka "Certyfikaty Użytkownika"
- Aplikuj o usługę "Globalny dostęp QosCosGrid"
 - zakładka "Moje Konto", ramka "Katalog Usług", "Usługi Globalne"



Recepta:

- Wygeneruj swój certyfikat gridowy
 - np. certyfikat z Simple CA na Portalu Użytkownika PL-Grid
 - zakładka "Moje Konto", ramka "Certyfikaty Użytkownika"
- Zainstaluj swój certyfikat gridowy w przeglądarce
- Przejdź na Portal Użytkownika PL-Grid - <https://portal.plgrid.pl>
- Zarejestruj swój certyfikat w Portalu PL-Grid
 - zakładka "Moje Konto", ramka "Certyfikaty Użytkownika"
- Aplikuj o usługę "Globalny dostęp gLite"
 - zakładka "Moje Konto", ramka "Katalog Usług", "Usługi Globalne"
 - rozprzestrzenianie informacji o użytkowniku w infrastrukturze może trwać do 6 godzin
- Aplikuj o usługę "InSilicoLab for Chemistry"
 - zakładka "Moje Konto", ramka "Katalog Usług", "Platforma dziedzinowa: Chemia"



Recepta:

- Wykonywanie obliczeń na kratach GPGPU wymaga dodatkowo dostępu do kilku dodatkowych usług. Dlatego:
- Aplikuj o usługę "Dostęp do UI - Cyfronet"
 - zakładka "Moje Konto", ramka "Katalog Usług", "Usługi dostępne"
- Aplikuj o usługę "Dostęp do GPGPU - Cyfronet"
 - zakładka "Moje Konto", ramka "Katalog Usług", " Usługi dostępne"
- Aplikuj o usługę "PLG-Data"
 - zakładka "Moje Konto", ramka "Katalog Usług", " Usługi dostępu do danych"
- Aplikuj o usługę "Rimrock"
 - zakładka "Moje Konto", ramka "Katalog Usług", " Usługi inne"



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.

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/>

- Dzień Otwarty ACK Cyfronet AGH: 23.11.2015
 - Rozstrzygnięcie konkursu na najlepszą pracę doktorską zrealizowaną w oparciu o zasoby ACK Cyfronet AGH





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

helpdesk@plgrid.pl

+48 12 632 33 55 wew. 312