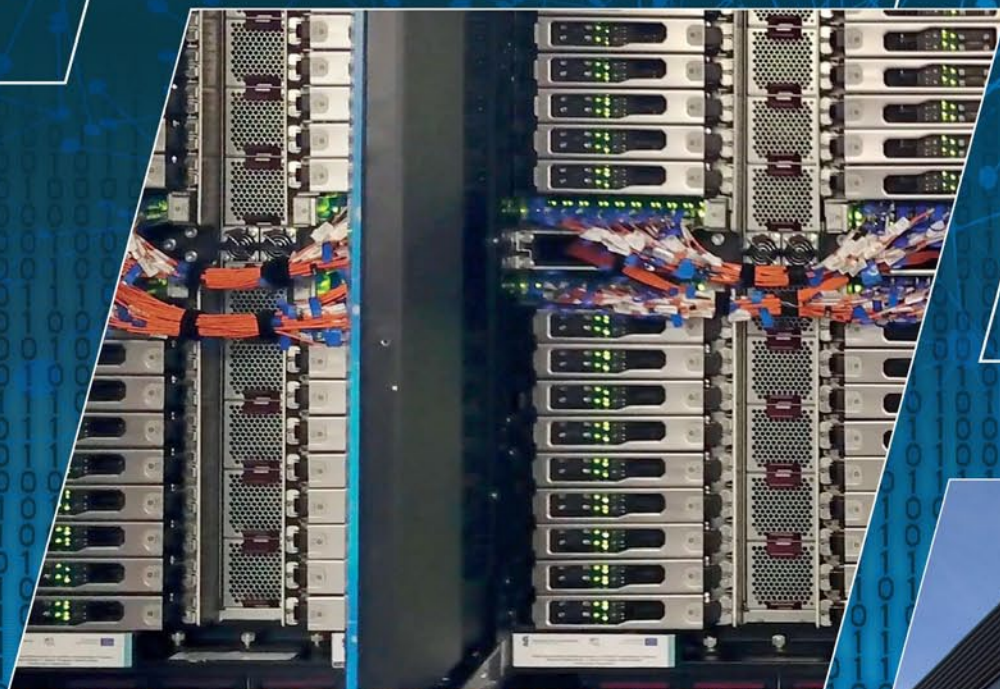




Akademickie Centrum Komputerowe
CYFRONET AGH

2022





ACK CYFRONET AGH jest jednostką wiodącą, ustanowioną przez Komitet Badań Naukowych, w zakresie eksploatacji i rozbudowy komputerów dużej mocy (KDM) oraz miejskiej akademickiej sieci komputerowej (MAN). CYFRONET jest koordynatorem Programu PLGrid oraz uznanym przez NCBiR Centrum Kompetencji w zakresie rozproszonych infrastruktur obliczeniowych typu gridowego i chmurowego.

Szanowni Państwo!

Za nami kolejny rok obciążony trudnościami pandemii Covid-19. Bogatsi jesteśmy o nowe doświadczenia i z nadzieją patrzymy w przyszłość, w nadchodzący kolejny rok. W roku 2022, oprócz realizacji licznych codziennych zadań, Cyfronet będzie kontynuował wiele prac związanych m.in. z udziałem w przedsięwzięciach: LUMI i EuroHPC PL.

LUMI to unikalna inicjatywa obejmująca dziesięć krajów europejskich, które w ramach Wspólnego Przedsięwzięcia EuroHPC (EuroHPC JU) inwestują w jeden system superkomputerowy o ogromnej szczytowej wydajności 552 PetaFlopsów. Dla Cyfronetu to duże wyzwanie, otrzymaliśmy bowiem zadanie zorganizowania dostępu dla polskich naukowców do jednego z najszybszych i najbardziej ekologicznych superkomputerów świata. LUMI ma przyspieszyć badania, a tym samym zwiększyć konkurencyjność rozwiązań pochodzących z Europy. Superkomputer będzie wspierał europejskich naukowców, przemysł i sektor publiczny w realizacji złożonych wyzwań m.in. w badaniach dotyczących medycyny, biologii, bioinformatyki, chemii, chemii kwantowej i fizyki molekularnej, fizyki wysokich energii, nowych materiałów, nanotechnologii, ekologii, energetyki, meteorologii i wielu wielu innych.



W ramach projektu „Narodowa Infrastruktura Superkomputerowa dla EuroHPC - EuroHPC PL” Cyfronet, jako koordynator projektu oraz lider Konsorcjum, buduje wraz z sześcioma Partnerami specjalistyczną infrastrukturę ogólnego przeznaczenia na potrzeby obliczeń wielkoskalowych, która umożliwi podejmowanie wyzwań badawczych w kluczowych z punktu widzenia polskiego społeczeństwa, środowiska naukowego i gospodarki obszarach. Infrastruktura ta czerpie z najnowocześniejszych technologii w zakresie High Performance Computing (HPC). Przeprowadzane na tej infrastrukturze obliczenia w wielkiej skali, obejmujące m.in.: symulacje oraz analizy dużych zbiorów danych naukowych, wpisują się w ramy przedsięwzięcia EuroHPC JU, tj. europejskiego systemu superkomputerów klasy eksaskalowej. Budowa i udostępnienie zasobów obliczeniowych w ramach EuroHPC PL znacząco wpłynie na zwiększenie zakresu i jakości badań naukowych opartych na danych oraz stosowania metod sztucznej inteligencji w Polsce.

W zestawieniu 500 najszybszych superkomputerów świata (lista TOP500, edycja z czerwca 2021 r.) obecne były cztery superkomputery z Polski, w tym dwa z Cyfronetu: Ares i Prometheus, zajmujące odpowiednio 216. i 373. miejsce. Uruchomiony w Cyfronecie w roku 2021 Ares jest zbudowany w ramach projektu PraceLab z zestawu serwerów obliczeniowych o mocy obliczeniowej ponad 3,5 PFlopsa oraz podsystemu dyskowego o pojemności ponad 11 PB. Superkomputer posiada 37 824 rdzenie obliczeniowe (procesory Intel Xeon Platinum) oraz 147,7 TB pamięci operacyjnej. Jest wyposażony w wydajne obliczeniowo karty GP GPU oraz system chłodzenia cieczą.

Tematyka badań naukowych realizowanych przy pomocy zasobów informatycznych Centrum dotyczy bardzo wielu dziedzin. Wspieramy naukowców korzystających z naszych superkomputerów w ich pracach naukowych i w dokonywaniu ważnych odkryć, a każdy sukces Użytkowników infrastruktury Cyfronetu jest dla nas niezmiernie istotny.

Bardzo dziękuję wszystkim Przyjaciółom i Użytkownikom Cyfronetu za współpracę i cenne wskazówki dotyczące dalszego rozwoju Centrum. Nieustannie zapraszam do osobistych kontaktów z pracownikami Cyfronetu. W obecnej sytuacji będą to nadal głównie kontakty zdalne, ale liczymy, że jak dotychczas przyniosą wiele owoców...

Z wyrazami szacunku

Prof. Kazimierz Wiatr

Dyrektor ACK Cyfronet AGH



DYREKTOR

prof. dr hab. inż. Kazimierz Wiatr

Tel. +48 12 633 34 26

e-mail: Kazimierz.Wiatr@cyfronet.pl



Zastępca Dyrektora

ds. Infrastruktury Informatycznej

mgr inż. Karol Krawentek

Tel. +48 12 633 34 26

e-mail: Karol.Krawentek@cyfronet.pl



Zastępca Dyrektora

ds. Komputerów Dużej Mocy

mgr Marek Magryś

Tel. +48 12 633 34 26

e-mail: Marek.Magrys@cyfronet.pl



Zastępca Dyrektora

ds. Administracyjnych

mgr Agnieszka Szymańska

Tel. +48 12 633 34 26

e-mail: Agnieszka.Szymanska@cyfronet.pl



Zastępca Dyrektora

ds. Ekonomiczno-Finansowych

mgr Angelika Zaleska-Walterbach

Tel. +48 12 633 80 53

e-mail: Angelika.Zaleska@cyfronet.pl

Dział Sieci Komputerowych – mgr Krzysztof Gawel (kierownik)

Tel. +48 12 634 10 25, e-mail: Krzysztof.Gawel@cyfronet.pl

Dział Komputerów Dużej Mocy – mgr inż. Patryk Lason (kierownik)

Tel. +48 12 632 33 55, e-mail: Patryk.Lason@cyfronet.pl

Dział Bezpieczeństwa Danych – dr inż. Grzegorz Sułkowski (kierownik)

Tel. +48 12 632 33 55, e-mail: Grzegorz.Sulkowski@cyfronet.pl

Dział Pamięci Masowych – mgr inż. Adrian Marszałik (kierownik)

Tel. +48 12 632 33 55, e-mail: Adrian.Marszalik@cyfronet.pl

Dział Oprogramowania HPC – mgr Łukasz Flis (kierownik)

Tel. +48 12 632 33 55, e-mail: Lukasz.Flis@cyfronet.pl

Dział Użytkowników – mgr inż. Robert Pająk (kierownik)

Tel. +48 12 632 33 55, e-mail: Robert.Pajak@cyfronet.pl

Dział Centrum Operacyjne – dr Andrzej Zemła (kierownik)

Tel. +48 12 632 33 55, e-mail: Andrzej.Zemla@cyfronet.pl

Dział Techniczny Nawojki – mgr inż. Damian Trela (kierownik)

Tel. +48 12 632 33 55, e-mail: Damian.Trela@cyfronet.pl

Dział Techniczny Pychowice – mgr inż. Mariusz Kula (kierownik)

Tel. +48 12 632 33 55, e-mail: Mariusz.Kula@cyfronet.pl

Dział Administracyjny – mgr Agnieszka Szymańska (kierownik)

Tel. +48 12 633 34 26, e-mail: Agnieszka.Szymanska@cyfronet.pl

Dział Ekonomiczno-Finansowy – mgr Angelika Zaleska-Walterbach (kierownik)

Tel. +48 12 633 80 53, e-mail: Angelika.Zaleska@cyfronet.pl

Dział Obsługi Projektów – mgr Aleksandra Pałuk (kierownik)

Tel. +48 12 632 33 55, e-mail: Aleksandra.Paluk@cyfronet.pl

Laboratorium Programu PLGrid – prof. Jacek Kitowski (kierownik)

Tel. +48 12 633 34 26, e-mail: kito@agh.edu.pl

Laboratorium Akceleracji Obliczeń i Sztucznej Inteligencji

dr hab. inż. Paweł Russek, prof. uczelni (kierownik)

Tel. +48 12 633 34 26, e-mail: Pawel.Russek@cyfronet.pl

Laboratorium Technologii Chmurowych – dr Łukasz Dutka (kierownik)

Tel. +48 12 633 34 26, e-mail: Lukasz.Dutka@cyfronet.pl

Laboratorium Przetwarzania Danych – mgr Tomasz Szepieniec (kierownik)

Tel. +48 12 632 33 55, e-mail: Tomasz.Szepieniec@cyfronet.pl

Laboratorium Obliczeń Kwantowych – dr Mariusz Sterzel (kierownik)

Tel. +48 12 632 33 55, e-mail: Mariusz.Sterzel@cyfronet.pl

Laboratorium Metod Informatycznych w Medycynie – dr inż. Marian Bubak (kierownik)

Tel. +48 12 633 34 26, e-mail: bubak@agh.edu.pl

Laboratorium Algorytmów Równoległych – prof. Bogusław Cyganek (kierownik)

Tel. +48 12 632 33 55, e-mail: cyganek@agh.edu.pl

Laboratorium Zastosowań Technik Obliczeniowych – dr hab. inż. Łukasz Rauch, prof. uczelni (kierownik)

Tel. +48 12 632 33 55, e-mail: lrauch@agh.edu.pl

Laboratorium Technik Wizualnych – mgr Jacek Przybylski (kierownik)

Tel. +48 12 632 33 55, e-mail: Jacek.Przybylski@cyfronet.pl



fot. Adam Frączek, KSAF AGH



fot. Krzysztof Mastalski, KSAF AGH

Kadencja 2021-2024

prof. dr hab. Marek Jarnicki
Przewodniczący

Uniwersytet Jagielloński

prof. dr hab. Tadeusz Lesiak
Zastępca Przewodniczącego

Instytut Fizyki Jądrowej PAN

mgr Jakub Badowski

Instytut Nafty i Gazu
Państwowy Instytut Badawczy

prof. dr hab. inż. Marek Gorgoń

Akademia Górniczo-Hutnicza

mgr inż. Jarosław Górski

Akademia Wychowania Fizycznego

dr Robert Gryboś

Instytut Katalizy i Fizykochemii Powierzchni PAN

prof. dr hab. Tomasz Hachaj

Uniwersytet Pedagogiczny

dr inż. Jerzy R. Jaworowski

Politechnika Krakowska

prof. dr hab. inż. Marek Kisiel-Dorohinicki

Akademia Górniczo-Hutnicza

mgr inż. Mariusz Kwinta-Pudęłko

Uniwersytet Rolniczy

mgr inż. Wojciech Majka

Uniwersytet Jagielloński Collegium Medicum

mgr inż. Tomasz Masier

Instytut Ekspertyz Sądowych

ks. dr Roman Mazur SDB

Uniwersytet Papieski Jana Pawła II

mgr Paweł Mikłaszewski

Akademia Muzyczna
Akademia Sztuk Pięknych
Akademia Sztuk Teatralnych

dr Wojciech Palacz

Uniwersytet Jagielloński

prof. dr hab. Dariusz Put

Uniwersytet Ekonomiczny

mgr Elżbieta Tomczyk

Sieć Badawcza Łukasiewicz
Krakowski Instytut Technologiczny

Z-ca Dyrektora Jacek Barski

Urząd Marszałkowski Województwa Małopolskiego

Dyrektor Paweł Schmidt

Urząd Miasta Krakowa

Ares – w kierunku krótszego czasu obliczeń

W roku 2021 został uruchomiony w Cyfronecie superkomputer Ares. Jest on zbudowany z serwerów obliczeniowych z procesorami firmy Intel (model Xeon Platinum), podzielonych na trzy grupy:

- 532 serwery wyposażone w 192 GB pamięci RAM,
- 256 serwerów wyposażonych w 384 GB pamięci RAM,
- 9 serwerów, każdy posiadający 8 kart NVIDIA Tesla V100.

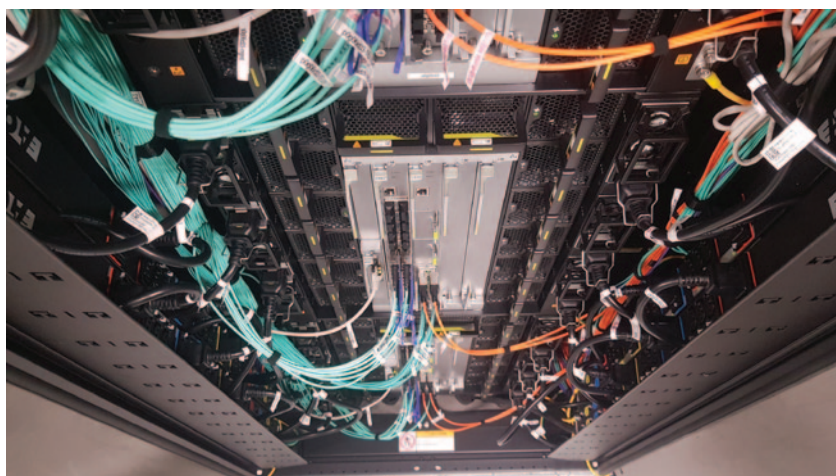
Ares w liczbach	
Liczba rdzeni obliczeniowych	37 824
Pamięć operacyjna	147,7 TB
Liczba kart z procesorami graficznymi GPGPU	72
Moc obliczeniowa	4 PFlops
Ranking TOP500 najszybszych komputerów świata (czerwiec 2021)	216 lokata

Sumaryczna, teoretyczna wydajność części CPU to ponad 3,5 PFlops, a części GPU to ponad 500 TFlops. Aresa wspomaga również system dyskowy o pojemności ponad 11 PB. Do przesyłania danych używana jest sieć typu InfiniBand EDR. Superkomputer posiada 37.824 rdzenie obliczeniowe oraz 147,7 TB pamięci RAM. Został też wyposażony w system chłodzenia cieczą.

Ares stanowi uzupełnienie zasobów obliczeniowych Cyfronetu poprzez dostarczenie nowszej generacji procesorów i serwerów z większą ilością

pamięci. Umożliwi to skrócenie czasu obliczeń zadań naukowych oraz zaadresowanie problemów, które do tej pory nie mogły być uruchamiane w dużej skali ze względu na niewystarczającą ilość pamięci. Dodatkowo, umiejscowienie Aresa w innej geograficznie lokalizacji niż Prometheus gwarantuje zachowanie ciągłości świadczenia usług obliczeniowych.





Prometheus – PetaFlops’owa Moc

Bardzo szybko następujące zmiany w świecie nauki wpływają na tempo rozwoju zaplecza informatycznego, oferowanego naukowcom przez Cyfronet. Rosnące coraz wyraźniej zapotrzebowanie środowisk naukowych na moc obliczeniową i coraz większe magazyny danych jest wyraźnie zauważalne ze strony dziedzin niemalże tradycyjnie związanych z Komputerami Dużej Mocy: chemii, fizyki, astronomii, szeroko rozumianych nauk o życiu oraz dziedzinach im pokrewnych. Astronomia, astrofizyka czy fizyka przestrzeni kosmicznej z jednej strony opierają się na akwizycji i analizie danych, a z drugiej na złożonych symulacjach komputerowych. Nauki biologiczne, chemiczne, medyczne, jak i wyżej wymienione, cechują się dynamicznym rozwojem oraz wprowadzaniem nowych, coraz bardziej wyrafinowanych metod badawczych, np. technik molekularnych opierających się na wysokowydajnym sekwencjonowaniu DNA. Multidyscyplinarna dziedzina, jaką jest medycyna, mierzy się z wieloma czasochłonnymi analizami, m.in. genomu ludzkiego. W efekcie wzrasta zapotrzebowanie na automatyczne zbieranie, przechowywanie i analizowanie sygnałów oraz obrazów biomedycznych, a do realizacji tych procesów niezbędne stają się zasoby superkomputerowe. Możliwość powiązania ze sobą wielu unikatowych danych, to jest danych klinicznych, genetycznych oraz środowiskowych i społecznych, niesie ze sobą wiele korzyści, ale i tu konieczne są dedykowane usługi, które mogą zaoferować jedynie centra superkomputerowe.

To właśnie z tego typu zadaniami mierzy się Prometheus – najszybszy polski superkomputer. Jako następca Zeusa, został włączony do infrastruktury PLGrid i nieodpłatnie służy naukowcom, także w ramach międzynarodowych projektów badawczych. Wy-

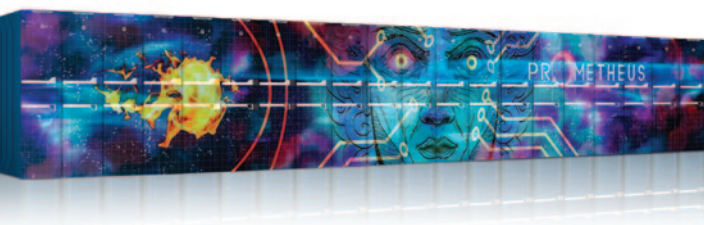
korzystanie superkomputera obejmuje: analizę wyników badań, symulacje numeryczne, przetwarzanie danych z sieci, zaawansowane wizualizacje oraz równoległą analizę ogromnych zbiorów danych (*big data*).

Prometheus składa się z ponad 2239 serwerów platformy HP Apollo 8000, połączonych superszybkością siecią InfiniBand FDR o przepustowości 56 Gbit/s. Superkomputer posiada 53 748 rdze-

ni obliczeniowych (procesorów Intel Haswell oraz Intel Skylake) oraz 283,5 TB pamięci operacyjnej w technologii DDR4. Został wyposażony w dwa systemy plików o łącznej pojemności 10 PB oraz bardzo dużej szybkości dostępu wynoszącej 180 GB/s. Prometheus jest również wyposażony w 144 karty NVIDIA z procesorami graficznymi GPGPU Tesla K40 XL oraz w 32 karty NVIDIA Tesla V100. Teoretyczna moc obliczeniowa superkomputera to 2,65 PFlops (PetaFlops)!

Co ważne, dzięki innowacyjnej technologii bezpośredniego chłodzenia cieczą procesorów i modułów pamięci operacyjnej, Prometheus zaliczany jest do najbardziej energooszczędnych systemów obliczeniowych tej klasy na świecie. Zostało to osiągnięte dzięki wykorzystaniu do chłodzenia mikroprocesorów cieczy o temperaturze 28°C, do wychłodzenia której w naszym klimacie wystarczą tzw. dry-coolery, zamiast konsumujących duże ilości energii elektrycznej generatorów wody lodowej oraz klasycznych klimatyzatorów technologicznych. Dzięki zastosowaniu chłodzenia cieczą, elementy elektroniczne pracują w niższych niż zwykle tempera-

turach, co pozytywnie wpływa nie tylko na ich mniejszą awaryjność, ale także pozwala uzyskać wydajność o ponad 5% większą niż dla analogicznej instalacji opartej na klasycznym chłodzeniu powietrzem. Ponadto, chłodzenie cieczą umożliwia osiągnięcie ekstremalnie wysokiej gęstości instalacji – 144 serwerów obliczeniowych, dzięki czemu ważąca ponad czterdzieści ton część obliczeniowa Prometheusa mieści się w zaledwie dwudziestu szafach. Ma to istotny wpływ także na wewnętrzną transmisję danych, bowiem odległości połączeń są tu krytyczne.



Cały system obliczeniowy, wraz z niezbędnymi elementami towarzyszącymi, w tym m.in. systemem gwarantowanego zasilania z dodatkowym agregatem prądotwórczym oraz nowoczesnymi systemami klimatyzacji technologicznej i gaszenia gazem, został zainstalowany w nowoczesnym Centrum Przetwarzania Danych (Data Center) Cyfronetu, specjalnie przystosowanym do eksploatacji Prometheusa.

Prometheus po raz kolejny znalazł się w zestawieniu 500 najszybszych komputerów świata (lista TOP500, edycja z czerwca 2021 r.) zajmując **373 miejsce, jako jeden z czterech superkomputerów z Polski**. Na liście po raz pierwszy pojawił się Ares, drugi superkomputer z Cyfronetu, który zajął 216 lokatę.

Ponieważ zastosowany w superkomputerze Zeus podział na partycje o różnej funkcjonalności został bardzo dobrze wykorzystany przez użytkowników, architektura superkomputera Prometheus składa się również z kilku klas węzłów o zróżnicowanym przeznaczeniu:

- klasycznej partycji ogólnego przeznaczenia z wydajnymi dwuprocesorowymi węzłami obliczeniowymi (Intel Xeon E5-2680v3 Haswell i Intel Xeon Gold 6128),
- zestawu serwerów wyposażonych w akceleratory graficzne GPGPU NVIDIA Tesla K40 XL,
- partycji akceleracyjnej wyposażonej w zestaw urządzeń uzupełniających konfigurację Prometheusa o kilka rodzajów akceleratorów (m.in.: karty GPGPU – NVIDIA K80, karty Intel Xeon Phi 7120P oraz karty Nallatech z układami FPGA Altera Stratix V: P395-AB2-330A-30 i P395-D8-330A-30),
- partycji dedykowanej dla obliczeń związanych ze sztuczną inteligencją, wyposażonej w akceleratory graficzne GPGPU NVIDIA Tesla V100. Warto wspomnieć, że **partycja ta stanowi system o mocy obliczeniowej ponad 4 PFlops dla operacji tensorowych i 256 TFlops dla standardowych obliczeń wykonywanych na liczbach podwójnej precyzji, co czyni go najszybszym dedykowanym rozwiązaniem dla sztucznej inteligencji dostępnym dla potrzeb nauki w Polsce.**

Prometheus w liczbach	
Liczba rdzeni obliczeniowych	53 748
Pamięć operacyjna	283,5 TB
Liczba kart z procesorami graficznymi GPGPU	144
Moc obliczeniowa	2,65 PFlops
Ranking TOP500 najszybszych komputerów świata (czerwiec 2021)	373 lokata



Rok	Liczba zadań	Czas CPU w latach
2015	1 099 822	5 811
2016	3 080 543	21 239
2017	5 032 438	36 600
2018	5 430 811	39 946
2019	2 738 534	41 829
2020	3 119 106	39 680

Dzięki uruchomieniu superkomputera Prometheus użytkownicy otrzymali ponad siedmiokrotnie większe możliwości w porównaniu z wcześniej wykorzystywanym Zeusem. Znacznie wydajniejsze procesory, szybsza sieć połączeń wewnętrznych oraz większa ilość pamięci operacyjnej pozwalają realizować obliczenia w skali niemożliwej do osiągnięcia na wcześniejszych zasobach Cyfronetu.

Zeus – ponad 60 000 lat obliczeń

Architektura superkomputera Zeus



Nieprzerwanie od roku 2008 w Cyfronecie działa superkomputer Zeus – długoletni lider wydajnościowy rankingów najszybszych systemów obliczeniowych w Polsce. Superkomputer ten posiada aktualnie 25 468 rdzeni obliczeniowych o łącznej teoretycznej mocy obliczeniowej 374 TFlops, 60 TB pamięci operacyjnej RAM oraz pamięć dyskową o pojemności 2,3 PB. Wyposażony dodatkowo w ponad 200 kart GPGPU, stanowi bardzo duże wsparcie dla obliczeń wykonywanych przez społeczności naukowe.

Od momentu instalacji Zeus nieprzerwanie notowany był na liście **TOP500**, najszybszych superkomputerów świata (12 razy), zajmując bardzo wysokie miejsca: cztery razy w pierwszej setce, a najwyżej – **na 81 pozycji**. Zeus był 10-krotnie najszybszy w Polsce.

Architektura

Zeus jest heterogenicznym superkomputerem obliczeniowym. Składa się z czterech klas węzłów, zróżnicowanych pod względem architektury zasobów obliczeniowych, specjalnie dostosowanych do wymagań środowisk naukowych:

- **klasycznego klastra** serwerów obliczeniowych, z wysoko wydajnymi węzłami CPU wyposażonymi w dwa procesory Intel Xeon i 16-24 GB pamięci na węzeł,
- klastra serwerów z dużą ilością pamięci – złożonego z tzw. „**grubych**” węzłów z czterema procesorami AMD Opteron i 256 GB pamięci na węzeł,
- zestawu serwerów wyposażonych w akceleratory **GPGPU** (procesory Intel Xeon i karty NVIDIA M2050 oraz karty NVIDIA M2090) i **FPGA** (moduły Pico Computing M-503 z układami Xilinx Virtex-6 LX240T),
- „**wirtualnego**” komputera **SMP** z dużą ilością pamięci współdzielonej, wykorzystującego dodatkowe oprogramowanie vSMP firmy ScaleMP, które pozwala na uruchamianie maszyn posiadających do 768 rdzeni i 6 TB pamięci RAM.

Zróżnicowanie rodzajów węzłów umożliwia dopasowanie aplikacji użytkowników do sprzętu, który jest najlepiej dobrany do cech i specjalnych wymagań tych aplikacji. Na przykład, grupa klasycznych węzłów CPU jest najczęściej wykorzystywana do realizacji szeregowych i równoległych

(MPI) zadań obliczeniowych, podczas gdy druga grupa („grube” węzły) jest idealna dla zadań wymagających dużej ilości pamięci operacyjnej. Węzły GPU pozwalają niektórym aplikacjom na korzystanie z akceleratorów GPGPU, a z kolei węzły vSMP umożliwiają uruchamianie zadań wymagających ogromnej pamięci lub pozwalają na skalowanie aplikacji, które nie korzystają z żadnej biblioteki komunikacji między węzłami, takiej jak MPI, w celu zrównoleglenia obliczeń.



Użytkownicy

Od początku swego istnienia, superkomputer Zeus służył całemu polskiemu środowisku naukowemu. W 2020 roku, tylko sam Zeus wykonał ponad **2,5 miliona zadań** dla naukowców z wielu uczelni i instytutów badawczych. Większość z nich realizowanych było na więcej niż jednym procesorze, a typowe obliczenia użytkownika żądają nawet **kilku tysięcy rdzeni dla jednego zadania!**

Rok	Liczba zadań	Czas CPU w latach
2008	603 525	207
2009	2 227 804	876
2010	4 009 049	990
2011	7 557 817	5 052
2012	8 126 522	7 923
2013	7 932 978	11 016
2014	7 694 224	12 980
2015	6 405 941	10 141
2016	4 668 134	3 414
2017	4 034 454	2 632
2018	2 911 875	2 490
2019	2 255 105	2 198
2020	2 577 813	2 081

Wykorzystanie superkomputerów

Prometheus i Zeus są częścią europejskiej chmurowej i gridowej infrastruktury w ramach Europejskiej Infrastruktury Gridowej (EGI). Jednocześnie, Prometheus i Zeus są również ważnymi superkomputerami w ogólnopolskiej infrastrukturze obliczeniowej PLGrid – platformie umożliwiającej prowadzenie badań naukowych *in silico* i wykonywanie obliczeń na komputerach dużej mocy, także w architekturze chmurowej i gridowej.

Dzięki infrastrukturze PLGrid możliwe jest intensywne wykorzystanie mocy obliczeniowych Prometheusa i Zeusa przez naukowców. Dedykowane środowiska obliczeniowe, tzw. gridy dziedzinowe, oraz specjalistyczne platformy informatyczne umożliwiają efektywną realizację coraz bardziej złożonych problemów obliczeniowych. Tematyka badań naukowych realizowanych przy pomocy superkomputerów ACK Cyfronet AGH jest bardzo bogata. Oto wybrane przykłady tematów obliczeń:

- analiza dynamiczna ceramicznych materiałów kompozytowych,
- modelowanie i optymalizacja prototypu energooszczędnego pojazdu napędzanego wodorowym ogniwem paliwowym,
- symulacje dynamiki molekularnej elektrolitów dla ogniw jonowych,
- użycie metod kwantowych do analizy obrazów,
- ekspresja genów u dorszy z Morza Bałtyckiego,
- badanie oddziaływania peptydów fuzyjnych wirusa grypy z błonami lipidowymi,
- analiza wpływu podziemnego zatłaczania płynów na spadki naprężeń indukowanych wstrząsów sejsmicznych,
- procesy mikrofizyczne w plazmie kosmicznej,
- przetwarzanie i rozpoznawanie mowy i pisma odręcznego.

Szeroki zakres tematów badawczych jest dowodem ciągle rosnącej liczby naukowców, którzy są świadomi korzyści płynących z wykorzystania superkomputerów takich jak Zeus czy Prometheus. Z pomocą tych superkomputerów można uzyskać wyniki złożonych symulacji wiele, wiele razy szybciej niż na zwykłym komputerze stacjonarnym. Obliczenia, które przy użyciu pojedynczych komputerów często zajęłyby wiele lat (w konkretnych rzeczywistych przypadkach ponad 150, 700 czy nawet 1000 lat), tu mogą być wykonane najczęściej w ciągu zaledwie kilku dni. Co ważne, użytkownicy Cyfronetu mogą korzystać z profesjonalnego wsparcia – od pełnej dokumentacji, poprzez szkolenia, aż po indywidualne konsultacje z ekspertami.

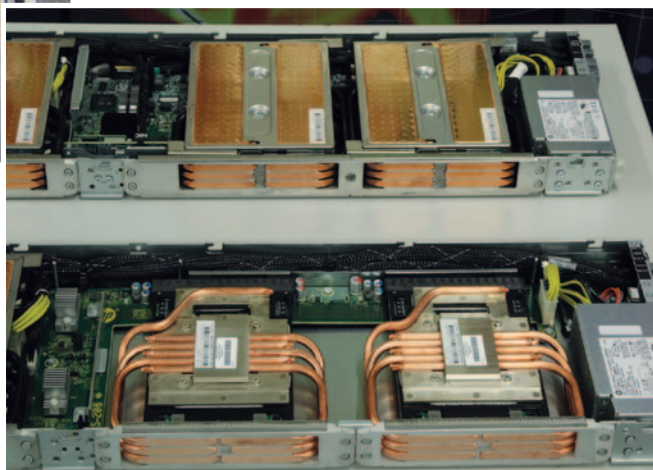
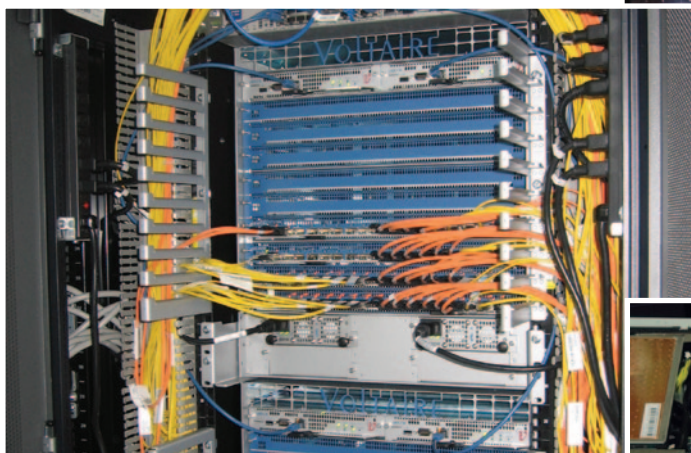
Poza indywidualnymi naukowcami i małymi grupami badawczymi obliczenia w ramach wielu różnych dyscyplin naukowych prowadzą z pomocą superkomputerów nawet międzynarodowe konsorcja, oczywiście z udziałem polskich naukowców. Dzięki superkomputerom w Cyfronecie polscy uczeni mogą uczestniczyć w ważnych projektach międzynarodowych, w tym w eksperymentach CTA, LOFAR, EPOS, Wielkiego Zderzacza Hadronów w CERNie oraz niedawno odkrytych fal grawitacyjnych w detektorach LIGO i VIRGO.

Oczywiście, nawet najwyższe lokaty w zestawieniu TOP500, czy też najnowocześniejsze technologie wykorzystane do budowy komputerów dużej mocy nie oddają w pełni znaczenia tego rodzaju zasobów obliczeniowych dla polskiego środowiska naukowego. O użyteczności superkomputerów udostępnianych przez ACK Cyfronet AGH jako narzędzia prowadzenia prac naukowo-badawczych, najlepiej świadczą dane statystyczne dotyczące ich wykorzystywania.

W tabeli przedstawione zostały zagregowane najważniejsze dane dotyczące liczby zadań obliczeniowych oraz czasu ich trwania, wykonanych przez Cyfronet dla innych jednostek.

Warto podkreślić, że ogromne zapotrzebowanie użytkowników na moc obliczeniową i przestrzeń do przechowywania danych nie zostałoby zaspokojone bez ciągłego powiększania zasobów obliczeniowych i pamięci dyskowej. Dlatego uważnie analizujemy sugestie użytkowników i dane statystyczne dotyczące prowadzonych obliczeń, nie zapominając jednocześnie o ciągłej obserwacji światowych trendów w informatyce.

Rok	Liczba zadań	Czas CPU w latach
Superkomputer Zeus		
2008	603 525	207
2009	2 227 804	876
2010	4 009 049	990
2011	7 557 817	5 052
2012	8 126 522	7 923
2013	7 932 978	11 016
2014	7 694 224	12 980
Superkomputer Prometheus		
2015	1 099 822	5 811
2016	3 080 543	21 239
2017	5 032 438	36 600
2018	5 430 811	39 946
2019	2 738 534	41 829
2020	3 119 106	39 680
Superkomputery Zeus i Prometheus łącznie		
2015	7 505 763	15 952
2016	7 748 677	24 653
2017	9 066 892	39 232
2018	8 342 686	42 436
2019	4 993 639	44 027
2020	5 696 919	41 761



Pamięć masowa dla superkomputerów

Odpowiednie skorelowanie projektu architektury infrastruktury obliczeniowej z właściwym doborem technologii rozwiązań składowania danych pozwala zoptymalizować usługi świadczone na rzecz użytkowników naukowych. Skala problemów w tym obszarze rośnie wraz ze złożonością i wydajnością eksploatowanych superkomputerów – obecnie systemy składowania danych dołączone do superkomputerów Cyfronetu przechowują ponad 500 000 000 plików o rozmiarach sięgających nawet kilku terabajtów. Szeroki zakres tematyczny prac badawczych realizowanych na zasobach udostępnianych przez Cyfronet znajduje swoje odbicie w różnorodności konfiguracji kluczowych superkomputerów Centrum, a przez to również w konfiguracji dedykowanych im zasobów pamięci masowej.

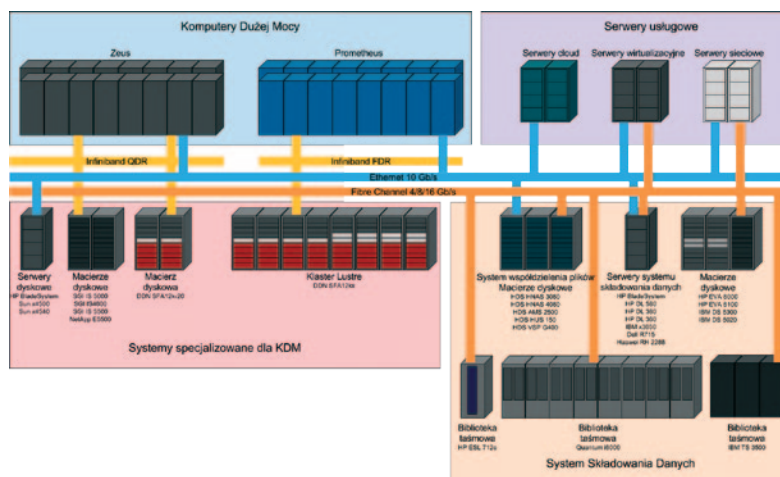
Najczęściej wykorzystywanym typem przestrzeni dyskowej są zasoby przeznaczone dla przechowywania katalogów domowych użytkowników. Przestrzeń tego rodzaju musi charakteryzować się bardzo wysokim poziomem dostępności oraz bezpieczeństwa danych, zwiększanego przez kopie migawkowe i zapasowe na zewnętrzne biblioteki taśmowe. Superkomputery Zeus i Prometheus oferują swoim użytkownikom uzyskanie opisanych funkcjonalności dzięki zastosowaniu specjalizowanych serwerów plików (tzw. głowic plikowych) HNAS. Serwery te wspierają sprzętowe implementacje protokołu NFS, zapewniając dużą wydajność, a także wysoką dostępność systemów plików. Głowice HNAS współpracują z macierzami dyskowymi Hitachi Data Systems AMS 2500 oraz HUS 150, służącymi jako repozytoria fizycznej przestrzeni dyskowej. Również i te urządzenia zapewniają niezwykle wysoki poziom bezpieczeństwa i wydajności, wynikający ze specyficznej charakterystyki danych przechowywanych w katalogach domowych.

W superkomputerach stosuje się także wydajną przestrzeń tymczasową „scratch”. Tu kluczowym elementem jest szybkość działania, dlatego jest ona oparta na bardzo szybkiej architekturze rozproszonego systemu plików – Lustre. Zaletą Lustre jest możliwość skalowania pojemności jak i wydajności przestrzeni dyskowej. W wyniku łączenia pojemności wielu serwerów, przepustowość I/O jest również agregowana i skaluje się z dodatkowymi serwerami. Ponadto, przepustowość i/lub pojemność może być łatwo zwiększona poprzez dynamiczne dodawanie większej liczby serwerów, bez konieczności przerywania obliczeń użytkowników. Obecnie w Cyfronie wszystkie superkomputery korzystają z przestrzeni „scratch” realizowanych przez Lustre. W przypadku Zeusa jest to system plików o pojemności niemal 600 TB i szybkości 12 GB/s, natomiast „scratch” Prometheusa charakteryzuje się pojemnością 5 PB i szybkością 120 GB/s, co czyni go najszybszym systemem plików w Polsce. Dodatkowo, dla zadań wymagających bardziej zaawansowanego dostępu do zasobów dyskowych, możliwe jest użycie super szybkiego dysku RAM, dostarczonego przez partycję vSMP superkomputera Zeus.

Większa część zasobów pamięci dyskowej Cyfronetu jest przeznaczona dla potrzeb użytkowników usług dziedzicznych rozwijanych w programie PLGrid. Infrastruktura PLGrid oferuje dedykowaną przestrzeń roboczą dla grup korzystających z usług dziedzicznych – funkcjonalność niezbędną do umożliwienia współpracy między naukowcami pracującymi w rozproszonych geograficznie lokalizacjach. Na Zeusie tego typu przestrzeń, o pojemności niemal 200 TB, dostarczana jest za pomocą głowic HNAS i protokołu NFS, natomiast Prometheus realizuje podobną funkcjonalność za pomocą systemu plików Lustre. Maksymalna pojemność zasobu */archive* w superkomputerze Prometheus to 5 PB, a sumaryczna szybkość operacji zapisu i odczytu sięga 60 GB/s.

Szczególnym przypadkiem pamięci masowej są zasoby dla dużych projektów i kolaboracji międzynarodowych, w których uczestniczy Cyfronet, takich jak WLCG (Worldwide LHC Computing Grid), analizujący dane z detektora LHC w CERN czy CTA (Cherenkov Telescope Array), badający promieniowanie gamma za pomocą sieci radioteleskopów. Tego typu projekty wymagają ogromnych zasobów dyskowych, dostępnych najczęściej za pomocą nietypowych protokołów takich jak SRM, xroot czy GridFTP. Cyfronet dostarcza tego typu przestrzeń dyskową przy użyciu kilku instancji dedykowanego oprogramowania DPM (Disk Pool Manager) i za pomocą dedykowanych sieci, takich jak LHCone. Łączna pojemność systemów DPM w Cyfronie przekracza 1 PB. Całkowita pojemność pamięci masowej przekracza **59 PB**.

Kompleksowa infrastruktura wydajnego i bezpiecznego składowania danych cyfrowych



Obserwowane obecnie zjawisko lawinowo rosnącej ilości wytwarzanych informacji w postaci cyfrowej dotyczy także środowiska naukowego. Dostęp do bardzo wydajnych superkomputerów umożliwia podejmowanie analiz wielkoskalowych problemów badawczych, w efekcie czego powstają olbrzymie zbiory danych. Wymagają one całkowicie nowego podejścia do procesów przetwarzania i przechowywania informacji. Problem ten, stanowiący obecnie jedno z najważniejszych wyzwań współczesnego świata cyfrowego, określany jest pojęciem BigData. Również i w ACK Cyfronet AGH wyraźnie widoczna jest korelacja pomiędzy wzrostem oczekiwań

w zakresie dostępnej pojemności, szybkości oraz dodatkowych funkcjonalności zasobów pamięci masowej, spowodowanym oferowaniem coraz wydajniejszych systemów obliczeniowych. Architektura Systemu Składowania Danych Cyfronetu – głównej platformy pamięci masowej dla Komputerów Dużej Mocy, złożona jest z kilku podstawowych elementów:

- sieci SAN – dedykowanej, wydajnej i wysokodostępnej sieci przeznaczonej do komunikacji urządzeń wchodzących w skład Systemu Składowania Danych oraz klientów wykorzystujących udostępniane zasoby lub usługi,
- macierzy oraz serwerów dyskowych oferujących przestrzeń przeznaczoną na składowanie danych użytkowników o dużej rozpiętości parametrycznej – począwszy od szybkich lecz równocześnie drogich i mało pojemnych rozwiązań, a skończywszy na urządzeniach bardzo pojemnych i relatywnie tanich, ale zarazem o ograniczonej wydajności,
- serwerów usługowych wraz ze specjalistycznym oprogramowaniem usługowym i wirtualizacyjnym, udostępniających dla użytkowników funkcjonalności takie jak: automatyczne kopie zapasowe i archiwalne, hierarchiczne systemy składowania danych, wydajne sprzętowe platformy plików bądź rozproszone sieciowe systemy plików,
- bibliotek taśmowych oraz specjalizowanego oprogramowania służącego do zapisywania krytycznych danych użytkowników na nośnikach magnetycznych,
- infrastruktury dodatkowej, w tym sieci Ethernet, Infiniband, rozwiązań wspomagających zarządzanie infrastrukturą IT oraz umożliwiających bezpieczne przechowywanie nośników magnetycznych.

Aktualnie całkowita fizyczna pojemność zasobów dyskowych Cyfronetu wynosi ponad **34 PB**.

Kopie zapasowe i archiwizacja

ACK Cyfronet AGH zapewnia swoim użytkownikom szerokie portfolio usług związanych z zabezpieczaniem informacji przechowywanej w postaci cyfrowej. Oprócz zaawansowanych rozwiązań technologicznych, takich jak: dedykowane systemom pamięci masowej sieci komunikacyjne, nowoczesne macierze dyskowe, czy też sprzętowe serwery plików, Centrum realizuje również usługi kopii zapasowych i archiwalnych, opartych o nośniki magnetyczne. Wbrew przewidywaniom mówiącym o nieuchronnym zmierzchu rozwiązań wykorzystujących składowanie danych na taśmach magnetycznych, technologia ta stale się rozwija, oferując w kolejnych generacjach nie tylko coraz większe pojemności nośników, lecz również zdecydowanie lepsze przepustowości oraz mechanizmy wspierające bezpieczeństwo i efektywność zapisu informacji (np. wbudowane w napędy taśmowe algorytmy szyfrujące i kompresujące dane).



Cyfronet dysponuje aktualnie trzema bibliotekami taśmowymi posiadającymi łącznie niemal 6 tysięcy slotów na taśmy magnetyczne standardu LTO oraz 36 napędów generacji III, IV, V i VI. Pojedynczy nośnik magnetyczny LTO-6 posiada pojemność fizyczną 2,5 TB i pozwala na zapis z prędkością sięgającą 160 MB/s, co teoretycznie umożliwia przechowywanie w bibliotekach taśmowych prawie 15 PB nieskompresowanych danych. Opisane zasoby wykorzystywane są do realizacji bieżących kopii zapasowych oraz archiwalnych istotnych zasobów informacyjnych użytkowników Centrum.

Kopia zapasowa wykonywana jest na aktywnych danych poprzez proces replikacji z lokalizacji źródłowej do wyodrębnionej i wyizolowanej lokalizacji docelowej. Procedura wykonania kopii zapasowej zapewnia spójność danych źródłowych i zapasowych, zarówno na poziomie pojedynczego pliku, jak i całego środowiska (co ma miejsce w przypadku złożonych systemów IT, takich jak serwery baz danych lub poczty oraz systemy wirtualne). Fizycznie, proces klonowania jest zwykle wykonywany przez kopiowanie danych źródłowych z dysku klienta na zasoby dyskowe / taśmowe docelowego serwera archiwizacyjnego, przy użyciu dedykowanego lub współdzielonego medium dostępowego, takiego jak Ethernet lub sieć SAN. Celem utworzenia archiwum jest natomiast zapewnienie bezpieczeństwa nieużywanych danych i zwolnienie zajmowanych zasobów pamięci masowej. W odróżnieniu od kopii zapasowej, archiwum jest tworzone tylko jeden raz, poprzez migrację danych z lokalizacji źródłowej do docelowej.

Cyfronet oferuje szeroki zakres usług do tworzenia kopii zapasowych, zarówno dostępnych bezpośrednio dla użytkowników, jak i działających automatycznie, bez ich ingerencji. Wśród tych do dyspozycji użytkowników, są usługi oparte na protokołach sieciowych FTP, NFS i SCP, działające w ramach dedykowanych serwerów kopii zapasowych. Udostępniają one rozwiązania służące do wykonywania kopii zapasowych, umożliwiając bezpośredni dostęp do zapasowych danych. W gestii użytkowników pozostaje natomiast decyzja, które dane mają pełnić rolę kopii zapasowych, a które archiwum.

W szczególnych przypadkach, Centrum oferuje usługę kopii zapasowych i archiwalnych, zwaną Systemem Powszechnej Archiwizacji. W ramach tej usługi, użytkownik otrzymuje dedykowaną przestrzeń dyskową, zabezpieczoną na wielu poziomach. Dane są składowane na macierzach dyskowych wyposażonych w mechanizmy nadmiarowości typu RAID 6, współpracujące z systemem hierarchicznego składowania danych HSM. W celu zapewnienia jeszcze wyższego poziomu bezpieczeństwa danych, kopie zapasowe danych użytkowników są dodatkowo chronione przez geograficzną replikację danych do jednostek stowarzyszonych. Użytkownicy Systemu Powszechnej Archiwizacji mogą także zwiększyć bezpieczeństwo swych danych poprzez ich szyfrowanie przy pomocy certyfikatów. Obecnie, całkowita pojemność zasobów taśmowych w Cyfronecie przekracza **25 PB**.

Program PLGrid
Infrastruktura i projekty



W ciągu ostatnich kilkunastu lat Cyfronet aktywnie uczestniczył w wielu krajowych i ogólnoeuropejskich projektach z zakresu technologii gridowych i cloudowych. Uzyskane w ten sposób wiedza i doświadczenie, a także podejmowane w wielu krajach Europy inicjatywy gridowe, przyczyniły się do sformułowania założeń i rozpoczęcia Programu PLGrid. Początkiem tego Programu było podpisanie porozumienia powołującego Konsorcjum PLGrid w 2007 roku, w skład którego weszły wszystkie akademickie centra superkomputerowe w Polsce: Interdyscyplinarne Centrum Modelowania Matematycznego i Komputerowego Uniwersytetu Warszawskiego, Poznańskie Centrum Superkomputerowo Sieciowe IChB PAN, Wrocławskie Centrum Sieciowo-Superkomputerowe PWR., Centrum Informatyczne Trójmiejskiej Akademickiej Sieci Komputerowej PG oraz **Akademickie Centrum Komputerowe Cyfronet AGH jako inicjator i koordynator Programu i Konsorcjum PLGrid**. Praca podjęta przez pięciu partnerów Konsorcjum PLGrid zaowocowała zbudowaniem niezawodnej, działającej 24 godziny na dobę, przez cały rok, rozproszonej infrastruktury obliczeniowej dla nauki, w ramach której udostępniana jest użytkownikom wielka moc obliczeniowa, a także duże zasoby pamięci masowej oraz dedykowane narzędzia umożliwiające tworzenie zaawansowanych aplikacji naukowych.



Proces budowy infrastruktury PLGrid rozpoczął się w trakcie projektu PL-Grid (Polska Infrastruktura Informatycznego Wspomagania Nauki w Europejskiej Przestrzeni Badawczej) i stanowił odpowiedź na rosnące potrzeby polskich środowisk naukowych, w których coraz istotniejszą rolę odgrywały zasoby obliczeniowe. Głównym celem zbudowanej infrastruktury było wsparcie badań naukowych poprzez integrację danych eksperymentalnych i wyników zaawansowanych obliczeń komputerowych, prowadzonych przez geograficznie rozproszone zespoły naukowców, przy użyciu zasobów centrów superkomputerowych. Cel ten został osiągnięty między innymi poprzez rozbudowę zasobów obliczeniowych we wszystkich instytucjach wchodzących w skład Konsorcjum. Co więcej, dzięki pracom realizowanym w projekcie PL-Grid, jesienią 2011 r. superkomputery wszystkich polskich akademickich centrów superkomputerowych znalazły się na TOP500 – liście najszybszych superkomputerów świata. Tego samego roku, superkomputer Zeus, zlokalizowany w Cyfronecie, znalazł się na 81. pozycji listy TOP500, będąc najszybszym superkomputerem w Polsce.

Kolejnym etapem realizacji Programu PLGrid było zapewnienie naukowcom profesjonalnego wsparcia informacyjnego poprzez opracowanie specyficznych środowisk obliczeniowych, czyli usług i oprogramowania oraz udzielanej pomocy użytkownikom w planowaniu, uruchamianiu i analizie wyników złożonych eksperymentów naukowych. Przygotowanie dedykowanych środowisk informatycznych, nazywanych gridami dziedzinowymi, dopasowanych do potrzeb 13 różnych grup naukowców, było najważniejszym zadaniem projektu PLGrid Plus (Dziedzinowo zorientowane usługi i zasoby infrastruktury PL-Grid dla wspomagania Polskiej Nauki w Europejskiej Przestrzeni Badawczej).

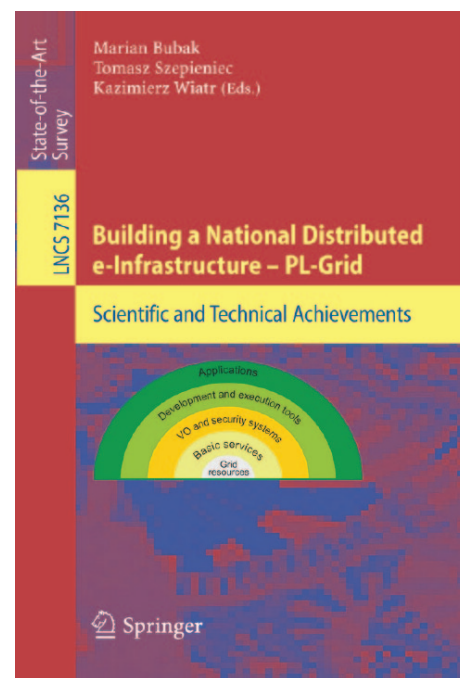
Przystosowanie infrastruktury do potrzeb naukowców konkretnych dziedzin było wielkim sukcesem projektu PLGrid Plus. Dlatego kierunek tych prac kontynuowano w ramach kolejnego projektu – PLGrid NG (Dziedzinowe Usługi Nowej Generacji w Infrastrukturze PL-Grid dla Polskiej Nauki), w ramach którego utworzone zostały usługi domeno-wo-specyficzne, spełniające wymagania kolejnych grup badawczych, reprezentujących 14 dziedzin nauki (**łącznie w obu projektach zbudowano narzędzia informatycznego wsparcia dla 27 dziedzin nauki**).

Uczestnicy Programu PLGrid nie przestali jednak na dostarczeniu domeno-wo-zorientowanych rozwiązań. Dzięki wieloletniemu zaangażowaniu w rozwój gridowych infrastruktur obliczeniowych, Cyfronet jest obecnie rozpoznawany jako Centrum Kompetencji w zakresie usług cloudowych i gridowych. Potwierdzeniem tego faktu było przyznanie ACK Cyfronet AGH środków na realizację nowego przedsięwzięcia: „Centrum kompetencji w zakresie rozproszonych infrastruktur typu gridowego – PLGrid Core”. Będąc kolejnym krokiem implementacji Programu PLGrid, projekt ten miał na celu rozwój możliwości infrastruktury, m.in. w zakresie obliczeń chmurowych i na dużych zasobach danych. Oprócz rozbudowy sprzętu komputerowego czy tworzenia nowego oprogramowania, prace w projekcie koncentrowały się także na infrastrukturze towarzyszącej. Jedną z ważnych inwestycji była budowa nowego zapasowego centrum danych, zlokalizowanego w pewnej odległości od dotychczasowej serwerowni Cyfronetu, co przełożyło się bezpośrednio na zwiększenie bezpieczeństwa przechowywanych danych naukowych.

W ramach prac prowadzonych w projekcie PLGrid Core, w Cyfronecie uruchomiono nowy **superkomputer Prometheus – w edycji listy TOP500 z listopada 2015 roku został on sklasyfikowany na 38 miejscu, najwyższym w historii polskich superkomputerów!**

Obecnie w ramach infrastruktury PLGrid udostępnianych jest ponad 5 PFlops mocy obliczeniowej i ponad 60 PB pamięci dyskowej. W infrastrukturze zaimplementowanych zostało wiele narzędzi wspomagających organizację skomplikowanych eksperymentów obliczeniowych, projektowanie i uruchamianie aplikacji, wspierających obliczeniowo badania naukowe oraz wizualizację otrzymywanych wyników. Udostępniono także nową usługę – obliczenia w „chmurze”.

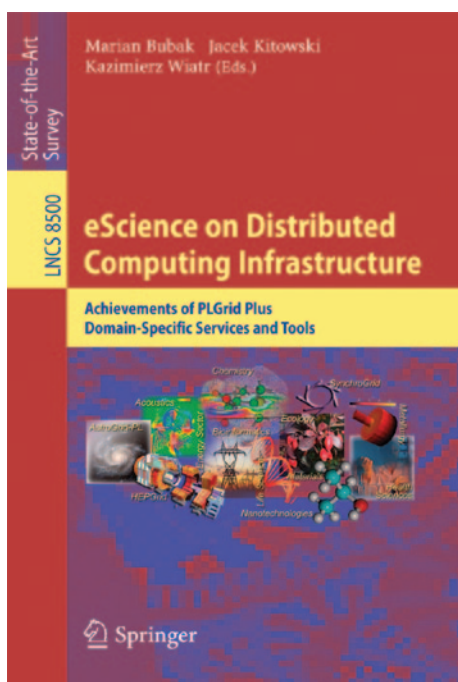
Wszystkie projekty realizujące cele Programu PLGrid uzyskały finansowanie m.in. ze środków Europejskiego Funduszu Rozwoju Regionalnego w ramach Programu Operacyjnego Innowacyjna Gospodarka. ACK Cyfronet AGH miał zaszczyt pełnić odpowiedzialną rolę koordynatora tychże projektów. Szeroki wachlarz oferowanych usług przyczynia się do umacniania współpracy pomiędzy polskimi i międzynarodowymi grupami specjalistów z wielu różnych dziedzin nauki – także humanistycznych i społecznych. Istotnym jest fakt, iż każda osoba prowadząca badania naukowe może być użytkownikiem infrastruktury. Dostęp do wielkich mocy obliczeniowych, ogromnych zasobów pamięci oraz zaawansowanych usług na światowym poziomie jest nieodpłatny dla polskich naukowców i wszystkich osób prowadzących działalność naukową, związaną z uczelnią lub instytutem naukowym w Polsce. Jedynym warunkiem jest założenie konta przez Portal PLGrid.



Usługi zorientowane dziedzinowo w infrastrukturze PLGrid



Infrastruktura PLGrid powstała dzięki inicjatywie Cyfronetu oferuje jednolity dostęp do zasobów wszystkich pięciu polskich centrów Komputerów Dużej Mocy. Unifikacja ma miejsce na wielu poziomach, poczynając od jednego loginu i hasła użytkownika w całej infrastrukturze, aż po dostęp do aplikacji naukowych.



Mając na uwadze sygnalizowaną przez użytkowników potrzebę implementacji dodatkowych mechanizmów wsparcia prowadzenia badań naukowych, infrastruktura PLGrid została rozszerzona o szereg specyficznych środowisk, rozwiązań i usług, opracowanych zgodnie z wymaganiami zdefiniowanymi przez 13 pilotażowych grup naukowców w ramach projektu PLGrid Plus (2011-2015). Głównym celem projektu było ułatwienie naukowcom korzystania z infrastruktury, a tym samym, przyciągnięcie nowych społeczności użytkowników, którzy potrzebują dużej mocy obliczeniowej i przestrzeni dyskowej superkomputerów, ale posiadają zbyt małe umiejętności w posługiwaniu się nimi. Aby umożliwić i ułatwić tworzenie środowisk domowo-specyficznych, w ramach projektu miała miejsce szeroka współpraca z przedstawicielami różnych dyscyplin, pracującymi często w konsorcjach dziedzinowych.

Dedykowane usługi dziedzinowe z jednej strony ukrywają złożoność infrastruktury, a z drugiej, w tym samym czasie, oferują funkcje istotne z punktu widzenia potrzeb

badaczy z danego obszaru tematycznego. W ten sposób użytkownicy otrzymują dokładnie taką funkcjonalność, jakiej potrzebują. Co więcej, jest ona im dostarczona w sposób specyficzny dla danej dziedziny, tak aby osiągnąć maksymalną intuicyjność i użyteczność.

Osiągnięcia naukowe i techniczne projektu PLGrid Plus zostały przedstawione w książce opublikowanej w wydawnictwie Springer, we wrześniu 2014 roku. Książka ta jest ważnym źródłem informacji dla naukowców, programistów i administratorów systemów, którzy w swojej pracy korzystają ze środowisk sieciowych i chmurowych. Książka zawiera 36 rozdziałów i jest podzielona na trzy części: pierwsza (rozdziały 1 do 8) zawiera ogólny przegląd prac wykonanych w ramach projektu i opis obecnego stanu infrastruktury PLGrid, w tym nowych rozwiązań w dziedzinie bezpieczeństwa i oprogramowania pośredniczącego. Druga część (rozdziały od

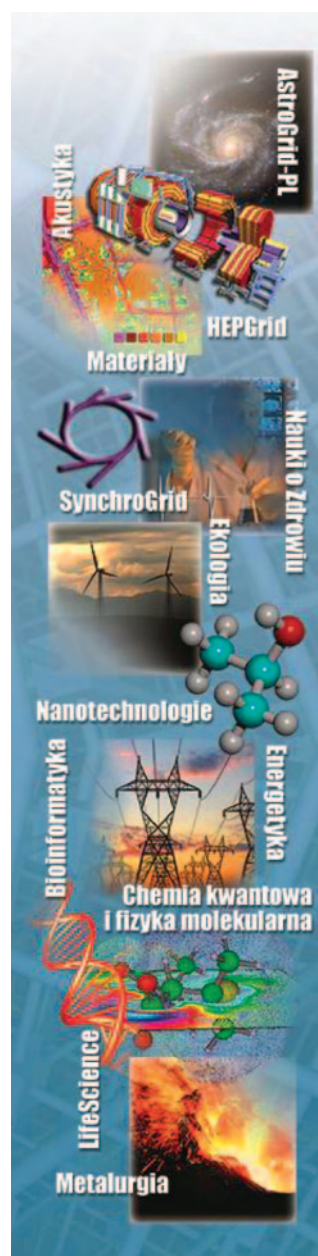
9 do 13) prezentuje nowe środowiska i usługi IT, które mogą być wykorzystywane przez wszystkie wyżej wymienione grupy naukowców. Trzecia część (rozdziały 14 do 36) opisuje, jak specyficzne środowiska, narzędzia i usługi, przygotowane w ramach projektu PLGrid Plus, są wykorzystywane w zaawansowanych obliczeniach i symulacjach komputerowych przeprowadzonych przez różne grupy naukowców. Rozdziały te przedstawiają modele obliczeniowe, nowe algorytmy i metody ich implementacji przy użyciu dostępnych narzędzi i usług.

Sukces projektu PLGrid Plus, a w szczególności rosnąca popularność specjalistycznych narzędzi i platform przygotowanych dla użytkowników z pierwszych 13 strategicznych obszarów nauki, doprowadziły do szybkiego wzrostu zapotrzebowania na podobne usługi dla naukowców z innych dziedzin. Dlatego, Konsorcjum PLGrid rozpoczęło projekt PLGrid NG (2014-2015), którego głównym celem było wdrożenie i udostępnienie, w ramach infrastruktury PLGrid, kilkunastu dodatkowych usług informatycznych dla grup naukowców reprezentujących 14 kolejnych obszarów nauki.

Nowe, zorientowane dziedzinowo usługi objęły szeroki zakres aktywności: począwszy od dostarczenia specjalistycznego oprogramowania, poprzez mechanizmy przechowywania danych, aż po nowoczesne platformy zintegrowane z nowym rodzajem narzędzi oraz dedykowanymi bazami danych, które przyspieszyły prowadzenie badań naukowych, jak również usprawniły i zautomatyzowały pracę grup badawczych.

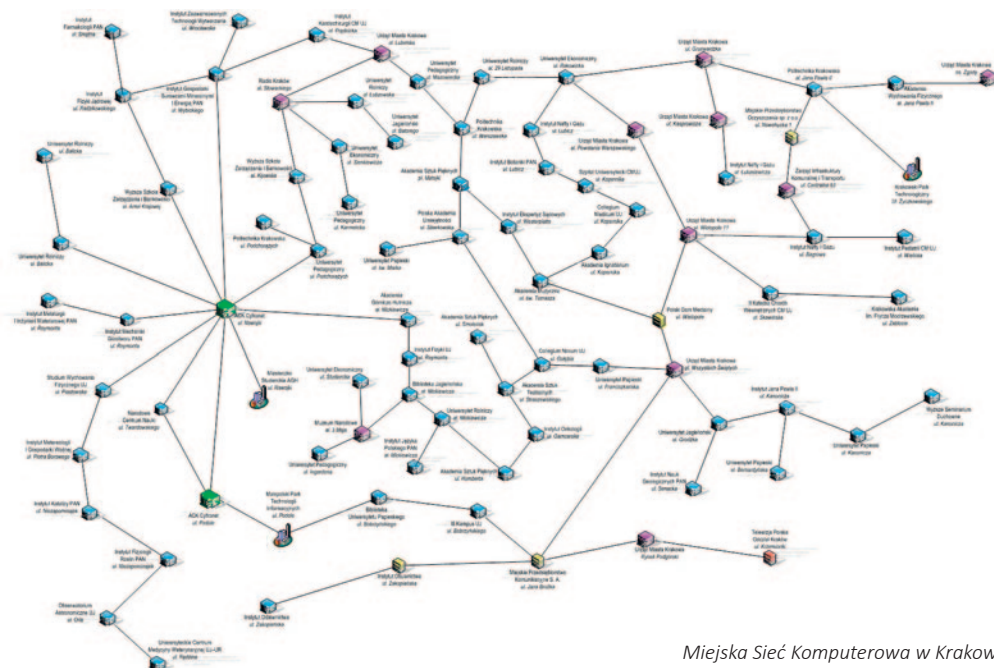
Przygotowanie i wdrożenie zestawu domenowo-specyficznych usług bardzo dobrze wpisało się w potrzeby wynikające z rozwoju zaawansowanej infrastruktury IT, przeznaczonej do realizacji nowoczesnych badań naukowych. Optymalnie dopasowana e-infrastruktura PLGrid spełnia nie tylko potrzeby naukowców dotyczące odpowiednich zasobów i usług obliczeniowych, ale także umożliwia polskim jednostkom naukowym współpracę z międzynarodowymi organizacjami badawczymi.

Rozbudowa istniejącej infrastruktury obliczeniowej w kierunku domenowo-specyficznych rozwiązań dla zespołów badawczych pozwoliła w efekcie końcowym na bardziej efektywne prowadzenie badań naukowych.



Miejska Sieć Komputerowa

ACK Cyfronet AGH stara się jak najlepiej służyć środowisku naukowemu, uczelniom i jednostkom naukowym w zakresie dostępu do krajowej i światowej sieci Internet oraz do mocy obliczeniowych na potrzeby badań naukowych. Dlatego ważnym zadaniem stawianym przed Cyfronetem jest zapewnienie bezawaryjnego funkcjonowania sieci oraz usług teleinformatycznych i obliczeniowych 7 dni w tygodniu przez 24 godziny na dobę.



Miejska Sieć Komputerowa w Krakowie

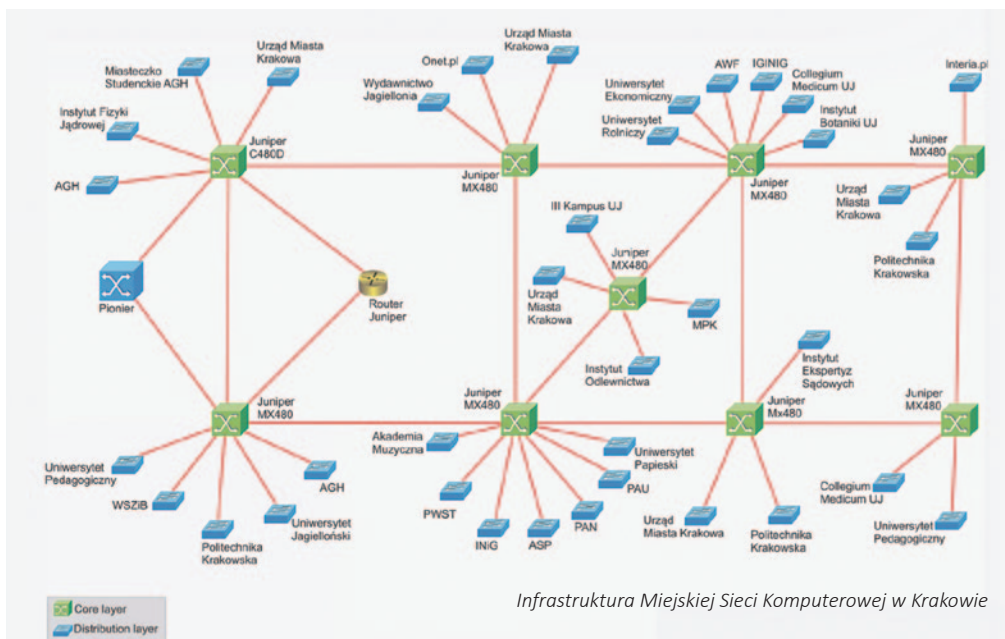
Charakterystyka MSK

ACK Cyfronet AGH, jako administrator Miejskiej Akademickiej Sieci Komputerowej MSK w Krakowie, stale rozbudowuje i modernizuje sieć, dostosowując ją do potrzeb i oczekiwań użytkowników. Obecnie, MSK eksploatuje własną infrastrukturę światłowodową o łącznej długości ok. 200 km. Światłowody MSK są ułożone w rejonie Starego Miasta, w okolicach kampusu AGH, docierają do Bronowic, Krowodrzy oraz do Czyżyn i Nowej Huty. Zrealizowane zostały także przyłączenia odległych instytutów w Prokocimiu, w rejonie Borku Fałęckiego oraz na terenie III Kampusu Uniwersytetu Jagiellońskiego w Pychowicach. Światłowody są układane w kanalizacji własnej, w dzierżawionej kanalizacji teletechnicznej oraz napowietrznie.

Infrastruktura światłowodowa jest podstawą funkcjonowania Miejskiej Akademickiej Sieci Komputerowej. ACK Cyfronet AGH dąży do tego, aby obejmowała ona jak największą liczbę obiektów uczelnianych i instytucji badawczych. Jednocześnie, z uwagi na stale rosnącą rolę

nowoczesnych środków łączności, w codziennej pracy bardzo ważnym jest, aby infrastruktura światłowodowa, oprócz dużej przepustowości, zapewniała również bezpieczną komunikację, uzyskiwaną w praktyce poprzez wykorzystanie łączy zapasowych pozwalających zachować ciągłość działania w sytuacjach zerwania tras podstawowych.

W warstwie rdzeniowej zastosowano nowoczesne przełączniki wyposażone w interfejsy 1 i 10 Gb Ethernet. Zapoczątkowano stopniowe doposażanie kluczowych przełączników w interfejsy 100 Gb. Wszystkie urządzenia warstwy rdzeniowej są połączone przynajmniej z dwoma lub trzema sąsiednimi, co pozwala automatycznie przełączyć ruch w przypadku awarii jednego z urządzeń lub awarii linii światłowodowej. Warstwę dystrybucyjną sieci oparto o przełączniki, z których każdy jest wyposażony w dwa interfejsy 1 Gb Ethernet (przyłączenie do warstwy rdzeniowej) oraz szereg interfejsów Ethernet typu 10/100/1000 Mbps służących do przyłączania użytkowników.



Miejska Akademska Sieć Komputerowa w Krakowie jest ważnym węzłem Akademickiej Sieci PIONIER – ogólnopolskiej szerokopasmowej sieci optycznej stanowiącej istotną bazę dla badań naukowych i prac rozwojowych w Polsce. MSK jest połączona w kierunku Warszawy, Katowic, Bielska-Białej i Rzeszowa łącami sieci PIONIER, każde o przepływności 2x10 Gbps. Łączność pomiędzy Centrami Komputerów Dużej Mocy w Polsce (Gdańsk, Kraków, Poznań, Warszawa i Wrocław) realizowana jest z przepustowością 2x100 Gbps. Za pośrednictwem sieci PIONIER realizowana jest komunikacja z wieloma ośrodkami krajowymi oraz zagranicznymi. Łączność zagraniczna odbywa się poprzez europejską naukową sieć GEANT o przepustowości 100 Gbps. Oprócz głównego połączenia do sieci GEANT, realizowane jest połączenie rezerwowe o przepustowości 5 Gbps do sieci CenturyLink Communications.

Usługi sieciowe dedykowane użytkownikom

Akademickie Centrum Komputerowe Cyfronet AGH od początku powstania polskiego Internetu (w połowie 1991 roku) aktywnie uczestniczy w rozwoju zarówno infrastruktury telekomunikacyjnej, jak też, co bardzo istotne, w rozwoju szeroko rozumianych usług internetowych. Realizowane na bardzo wydajnych komputerach usługi obejmują między innymi:



- **e-mail** – internetowa poczta elektroniczna, z dostępem poprzez protokół SMTP lub interfejs webowy <http://poczta.cyfronet.pl>,
- **www** – udostępniane serwisy, obok nowości ze świata nauki, prezentują informacje dotyczące kultury, zabytków oraz wielu innych dziedzin,
- **news** – forum dyskusyjne podzielone na grupy tematyczne z wielu dziedzin, w tym dyskusje naukowe przeznaczone dla wąskiego grona specjalistów,
- **dns** – serwery systemu nazw domen – obsługujące usługę tłumaczenia nazw sieciowych na adres IP dla użytkowników Miejskiej Akademickiej Sieci Komputerowej,
- **ftp** – serwer przechowuje i stale aktualizuje kopie zagranicznych archiwów oprogramowania (tzw. *mirror*), a także zawiera oprogramowanie typu shareware i freeware pracujące pod systemami MS-Windows oraz UNIX; dzięki uruchomieniu serwera ftp zostało znacznie ograniczone obciążenie łącza międzynarodowego, a także zwiększono wygodę użytkowników Miejskiej Akademickiej Sieci Komputerowej,
- **eduroam** – umożliwia łączność z siecią akademicką w każdym miejscu na świecie objętym zasięgiem eduroamu, przy użyciu jednego autoryzowanego konta,
- **box** – dysk sieciowy (<http://box.cyfronet.pl>) umożliwiający wymianę plików i ich synchronizację; dzięki dedykowanym aplikacjom może być również używany na urządzeniach przenośnych.

Usługi sieciowe w liczbach (w roku 2020)

Liczba wysłanych i odebranych wiadomości e-mail	~ 17 000 000
Liczba sesji nawiązanych z serwerem pocztowym	~ 45 000 000

Portale i aplikacje mobilne

Centrum nie ogranicza swojej działalności jedynie do obszarów powiązanych z nauką, ale przyczynia się także do rozwoju społeczeństwa informacyjnego. Serwer WWW w ACK Cyfynet AGH pełni funkcję węzła internetowego dla całej krakowskiej społeczności naukowej. Centrum wciąż ulepsza i rozbudowuje swój portal internetowy, który na przestrzeni ostatnich lat zyskał dużą popularność. W roku 2021 została opublikowana nowa wersja portalu, wraz z wersją dla urządzeń mobilnych.



Szczególne znaczenie dla Centrum ma współpraca z władzami Krakowa. Umowa pomiędzy Gminą Kraków i Cyfynetem, dotycząca promocji miasta, zaowocowała stworzeniem portalu internetowego, który oprócz prezentowania informacji naukowych zapoznaje czytelników z kulturą, zabytkami, turystyką, transportem lokalnym i wieloma innymi aspektami życia w Krakowie.

We współpracy z Urzędem Miasta Krakowa został przygotowany i uruchomiony Biuletyn Informacji Publicznej dla regionu krakowskiego. W roku 2005, współpraca ta została rozszerzona o świadczenie usług w zakresie publikacji treści dla miejskich instytucji, bibliotek, szkół, itp.

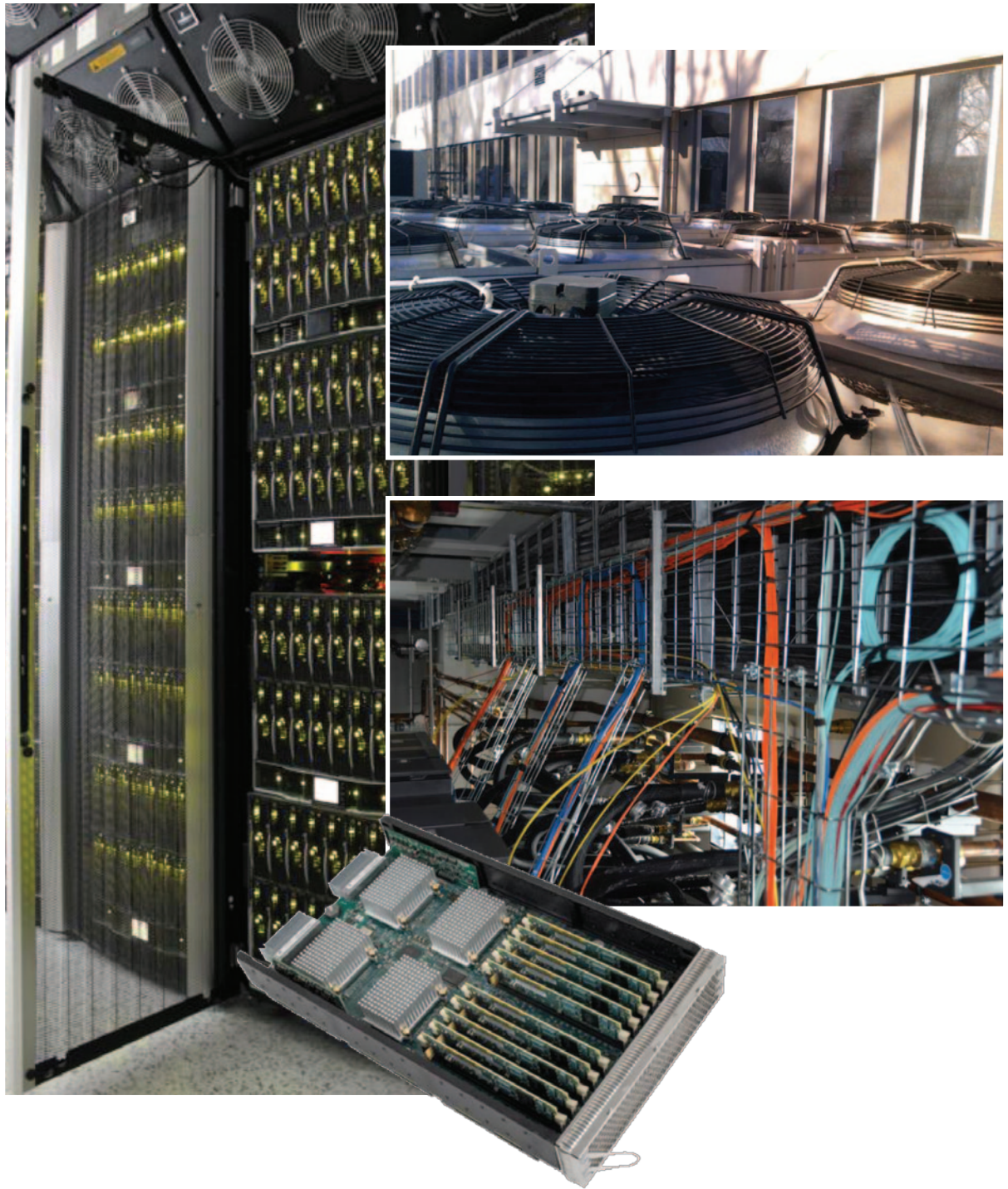
Najnowsza wersja oficjalnego serwisu internetowego miasta Krakowa została opublikowana jako Miejska Platforma Internetowa „Magiczny Kraków”. W roku 2007, portal ten – dostępny pod adresem www.krakow.pl – został



nominowany do World Summit Award jako najlepszy serwis e-administracji w Polsce. Mobilna wersja portalu została natomiast nagrodzona podczas konferencji Mobile Trends w roku 2012 jako najlepsza mobilna strona miejska w Polsce.

Współpraca z Urzędem Miasta wchodzi również w obszar urządzeń mobilnych – Cyfynet opracował m.in. mobilną aplikację *Kraków.pl*, która może być wykorzystana jako przewodnik miasta Krakowa i źródło ważnych informacji, takich jak numery telefonów, lokalizacje punktów informacyjnych, konsulatów lub aptek. Co ważne, aplikacja umożliwia znalezienie wszystkich tego typu miejsc w trybie mapy off-line. Aplikacja *Kraków.pl* jest dostępna w językach polskim, angielskim i hiszpańskim.



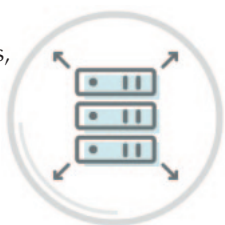


Zasoby obliczeniowe

ACK Cyfronet AGH dysponuje dojrzałą i unikatową infrastrukturą obliczeniową dostępną dla polskich naukowców. Jest ona oparta na pięciu filarach, rozbudowanych o szereg usług oraz rozwiązań przygotowanych wraz z naukowcami na potrzeby badaczy z różnych dziedzin nauki. Wraz ze szkoleniami i kompleksowym wsparciem, zasoby dostępne w naszym centrum obliczeniowym pozwalają na efektywne przeprowadzanie wszelkich badań naukowych wykorzystujących infrastrukturę komputerową.

Zasoby obliczeniowe

Superkomputery Prometheus, Zeus i Ares dostarczają:
7+ PFlops
117 000+ rdzeni
450+ GPGPU
500+ TB RAM



Zasoby dyskowe

34 PB przestrzeni dyskowej i 25 PB zasobów taśmowych na dane wraz z szybkim systemem plików Lustre umożliwia analizę wielkich zbiorów danych.



Profesjonalne oprogramowanie

Narzędzia i programy oraz biblioteki, które umożliwiają efektywne wykorzystanie zasobów obliczeniowych Cyfronetu, a także usprawniają pracę w różnych dziedzinach nauki.



Narzędzia pracy zespołowej

Serwisy i narzędzia takie jak serwer repozytoriów Stash Git oraz środowisko do zarządzania projektami JIRA ułatwiają realizację i koordynację złożonych projektów oraz współpracę zespołów badawczych.



Chmura obliczeniowa

Elastyczne rozwiązanie oparte o platformę OpenStack, pozwalające na łatwe dostosowanie środowiska obliczeniowego do potrzeb grupy badawczej.

Zaawansowane platformy obliczeniowe oraz usługi dla domen badawczych

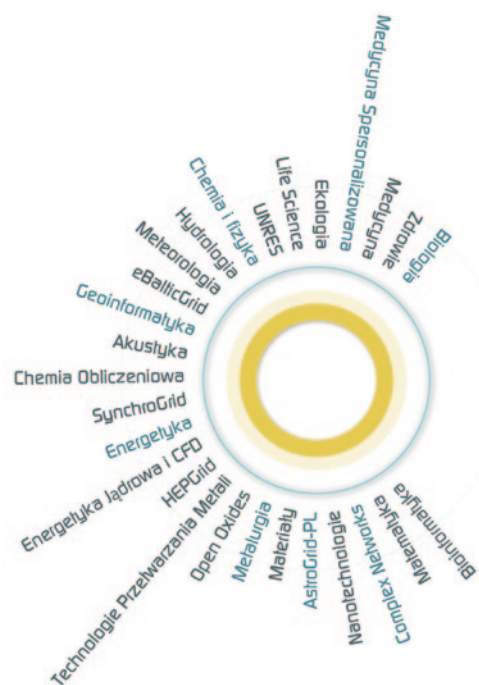
Naukowcy prowadzący prace badawcze z wykorzystaniem superkomputerów i dużych zasobów pamięci masowej mają różne wymagania oraz sposoby pracy, zatem potrzebują różnego typu dostępu do komputerów oraz innych zasobów infrastruktury IT. Odpowiadając na potrzeby użytkowników, Cyfronet udostępnia szereg zaawansowanych platform informatycznych i dedykowanych usług. Ukrywają one złożoność infrastruktury i jednocześnie dostarczają funkcjonalności istotne z punktu widzenia badaczy z danej dziedziny, precyzyjnie dopasowane do ich potrzeb.

Wraz z infrastrukturą obliczeniową udostępniamy zbiór narzędzi, który umożliwia badaczom wykonywanie złożonych eksperymentów o wielkiej skali oraz efektywne i łatwe zarządzanie ich rezultatami. Efektywność obliczeń i analiz oraz bezpieczeństwo analizowanych danych są gwarantowane poprzez odpowiedni dobór rozwiązań IT rozwijanych oraz konfigurowanych przez doświadczonych programistów zatrudnionych w Cyfronecie. Platformy prezentowane poniżej zostały z powodzeniem zastosowane w gridach dziedzinowych dostępnych w Programie PLGrid. Wszystkie narzędzia, usługi oraz serwisy są udostępniane w ramach infrastruktury PLGrid, co pozwala polskim naukowcom i ich zagranicznym współpracownikom na dostęp do zasobów oraz danych w najbardziej dla nich wygodny sposób.

W Centrum oferujemy m.in. zaawansowane narzędzia i interfejsy graficzne umożliwiające budowę dedykowanych środowisk do prowadzenia badań naukowych, przeprowadzanie wirtualnych eksperymentów, budowę portali webaplikacji, wizualizację wyników obliczeń, uruchamianie złożonych scenariuszy obliczeniowych, jak również wspierające jednolity i efektywny dostęp do danych. Wszystkie te usługi są ważnym wsparciem dla osób prowadzących badania naukowe, wpływają bowiem na usprawnienie i, tam gdzie to możliwe, zautomatyzowanie pracy grup badawczych, co znacznie przyspiesza uzyskanie wyników badań naukowych. Na kolejnych stronach znajdują się opisy ich możliwości.

Zaproszenie do współpracy

Poszukujemy osób zainteresowanych rozwojem usług obliczeniowych, w tym tych zorientowanych na konkretne dziedziny badawcze. Oferujemy również wsparcie w prowadzeniu badań naukowych. Zachęcamy naukowców do udostępnienia swoich autorskich kodów ekspertom naszego Centrum, którzy będą mogli pomóc w ich efektywnej kompilacji na naszych superkomputerach. Równocześnie zapewniamy pomoc w ich użytkowaniu i efektywnym wykorzystaniu. Umożliwiamy również wykorzystanie własnych licencji oprogramowania naukowego posiadanych przez grupy badawcze, gdy umożliwiają one obliczenia na superkomputerach Cyfronetu.



Chmura obliczeniowa w infrastrukturze PLGrid



W trakcie tworzenia infrastruktury PLGrid w znacznej mierze kierowano się potrzebami naukowców. Dlatego parametry infrastruktury pozwalają na jej łatwe dostosowanie do rozwiązywania nawet skomplikowanych problemów badawczych przez różne grupy badaczy – począwszy od małych zespołów, a skończywszy na międzynarodowych konsorcjach naukowych. Aby spełnić ich wymagania, obok typowych platform obliczeniowych i przechowywania danych, oferujemy „chmurę obliczeniową”.



– Do tej pory, infrastruktura PLGrid dostarczała szeregu platform informatycznych, wraz z zasobami obliczeniowymi i pamięcią masową. Chmura obliczeniowa nie stanowi jedynie ich rozszerzenia. Traktujemy ją jako nowy poziom jakości w prowadzeniu badań – mówi Kazimierz Wiatr, Dyrektor Centrum. – Użytkownik może w łatwy sposób połączyć się z żądanym zestawem maszyn wirtualnych (VM), z pełnymi prawami dostępu do systemu operacyjnego. W celu zapewnienia wysokiego poziomu bezpieczeństwa, wszystkie maszyny wirtualne działają w dedykowanej sieci lokalnej. Dostęp do poszczególnych usług można uzyskać z całego świata, co ułatwia współpracę między naukowcami – dodaje Dyrektor.

Należy podkreślić następujące zalety płynące z korzystania z chmury obliczeniowej:

- Chmura zwiększa elastyczność badań, ponieważ naukowcy mogą dostroić maszyny wirtualne do swych specyficznych potrzeb. Do tej pory, aby skonfigurować „wirtualne laboratorium” służące do rozwiązania określonego problemu naukowego, użytkownik potrzebował pomocy ze strony ekspertów PLGrid. Obecnie, każdy naukowiec może samodzielnie stworzyć i łatwo rozbudować takie wirtualne laboratorium.
- Katalog maszyn wirtualnych oferowanych przez PLGrid wspiera wiele systemów operacyjnych. Dzięki temu, użytkownicy mogą uruchamiać swoje aplikacje w systemach operacyjnych innych niż Scientific Linux, np. w Windows lub innych systemach linuxowych.
- Chmura obliczeniowa pozwala na łatwe zbudowanie i uruchomienie środowiska testowego. Ta funkcjonalność jest bardzo przydatna dla naukowców tworzących swoje własne oprogramowanie. Mogą w prosty sposób uruchomić każde zadanie testowe i szybko przeanalizować jego wyniki.
- Możliwe jest utrzymanie komunikacji z już uruchomionym zadaniem obliczeniowym. Ponadto, każda maszyna wirtualna może być łatwo powielana, nawet w tysiącach lub więcej egzemplarzy. Uruchomienie nowej maszyny wirtualnej zajmuje niespełna 30 sekund.
- Chmura obliczeniowa jest także najlepszym, a w wielu przypadkach jedynym, rozwiązaniem do uruchamiania zadań za pośrednictwem starszych pakietów oprogramowania. W bezpiecznym środowisku sieci LAN można korzystać nawet z przestarzałych, wycofanych już z powszechnego użycia, systemów operacyjnych. Ta funkcjonalność jest również dobrym rozwiązaniem dla rozproszonych grup międzynarodowych, wykorzystujących w swych badaniach wiele różnych pakietów. Każda grupa może prowadzić własne obliczenia i w prosty sposób udostępniać swoje wyniki innym.

– Chmura obliczeniowa w PLGrid i Cyfronet to innowacyjne rozwiązanie na skalę europejską. Mamy silne przekonanie, że wprowadzi ona nowy poziom jakości dla badań prowadzonych przez naszych użytkowników – podsumowuje Profesor Wiatr.

Obecnie, w ramach zasobów Cyfronetu eksploatowanych jest **ponad 200 różnych typów maszyn wirtualnych**.

ONEDATA

Onedata jest globalnie skalowalnym systemem zarządzania danymi, który pozwala na przezroczysty dostęp do danych przechowywanych w rozproszonych systemach składowania danych. Onedata nadaje się równie dobrze dla małych grup użytkowników jak i dla dużych międzynarodowych środowisk naukowych, dzięki czemu mogą korzystać z tego systemu zarówno federacje użytkowników naukowych jak i komercyjnych. Onedata umożliwia użytkownikom korzystanie z jednorodnego systemu zarządzania danymi zarówno do przechowywania danych osobistych jak i związanych z pracą zawodową, np. wyników badań, oraz daje możliwość dostępu do nich w wydajny sposób z dowolnego urządzenia.



Onedata pozwala na budowanie federacji zaufanych użytkowników oraz dostawców przestrzeni dyskowej, co pozwala użytkownikom z jednej strony na korzystanie z rozproszonej przestrzeni dyskowej a z drugiej na kontrolowanie ilości zasobów udostępnianych poszczególnym użytkownikom lub grupom roboczym przez administratorów, bez konieczności stosowania certyfikatów.

Zalety dla użytkowników

- Zunifikowany dostęp do danych rozproszonych pomiędzy dostawców przestrzeni dyskowej opartych o heterogeniczne systemy. Korzystając z Onedata, użytkownicy mają dostęp do swoich danych w każdym miejscu i z dowolnego urządzenia.
- Dane użytkowników zorganizowane są w wirtualne foldery (ang. *Space*), które mogą być rozproszone pomiędzy dostawców przestrzeni dyskowej i dostępne przez użytkowników z przeglądarki WWW lub poprzez protokół POSIX.
- Łatwy w użyciu Graficzny Interfejs Użytkownika pozwalający na dostęp i wyszukiwanie danych oraz zarządzanie nimi.
- Wsparcie dla łatwego dzielenia się i współpracy z innymi użytkownikami w oparciu

o dane, przy zachowaniu wysokiego poziomu bezpieczeństwa opartego o listy kontroli dostępu oraz złożony system grup i uprawnień do nich.

- Funkcjonalność publikowania i dostępu do danych otwartych oraz możliwość łatwej rejestracji identyfikatorów typu DOI dla zbiorów danych.

Zalety dla administratorów

- Łatwość instalacji i konfiguracji poprzez kontenery Docker i klientów powłoki.
- Uproszczony mechanizm przy zachowaniu pełnej kontroli administratorów w oparciu o system tokenów.
- Złożony system monitorowania dostarczający informacji o wszystkich aspektach systemu, dostępny poprzez interfejs REST oraz w postaci graficznej w panelu administracyjnym.
- Wsparcie dla wielu systemów składowania danych, w tym POSIX (np. Lustre), Amazon S3, Ceph, OpenStack Swift, czy GlusterFS.

Zalety dla twórców aplikacji

- Łatwa integracja z systemem poprzez interfejsy REST i CDML.
- Elastyczna autentykacja i autoryzacja dostępna o technologie OpenID Connect i Macaroons.
- Kompletna dokumentacja interfejsów w postaci pozwalającej na automatyczne generowanie kodu aplikacji klienckich odpowiedzialnego za komunikację z usługami Onedata.



Użytkownicy Onedata

Onedata jest aktualnie uruchomione oraz testowane w ramach kilku infrastruktur europejskich, w tym w projektach: PLGrid, INDIGO-DataCloud, EGI DataHub, Open Nebula Science Cloud oraz Human Brain Project. W projekcie HBP Onedata wykazało możliwość dostarczenia systemu zarządzania danymi pozwalającego na spełnienie wymagań przepustowości dla aplikacji wizualizacji przekrojów mózgu w czasie rzeczywistym.

Więcej informacji można znaleźć pod adresem <https://onedata.org>.

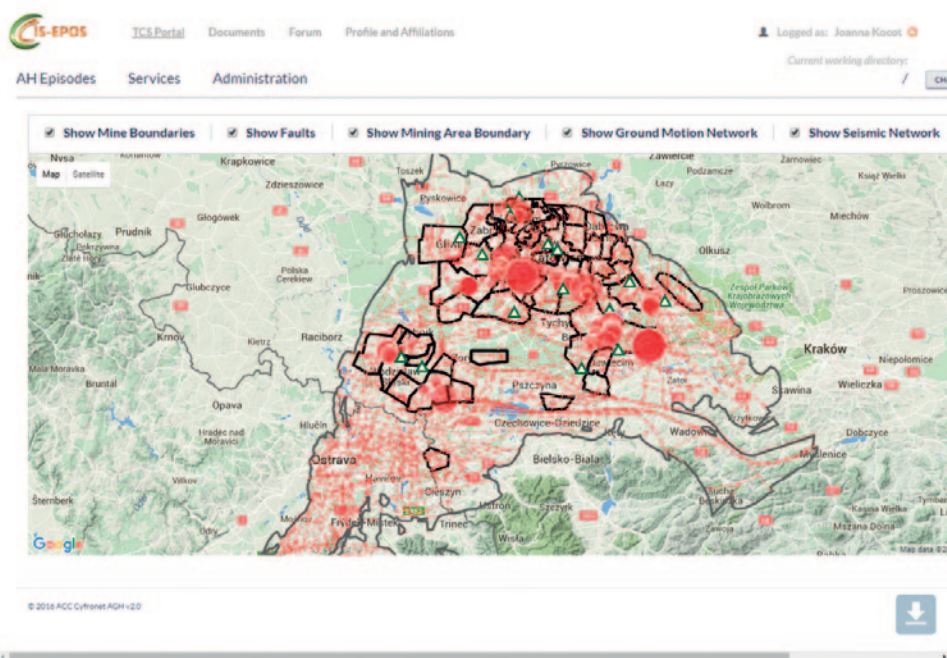


InSilicoLab to zestaw zaawansowanych narzędzi i bibliotek programistycznych pozwalających opracowywać i rozwijać portale oferujące środowiska pracy dla naukowców. Podstawowym celem stojącym za budową takiego środowiska programistycznego jest możliwość tworzenia dedykowanych portali naukowych, które z jednej strony pozwalają w pełni wykorzystać możliwości obliczeniowe rozproszonej infrastruktury, z drugiej zaś pozwalają użytkownikom prowadzić eksperymenty *in silico* w sposób odzwierciedlający ich typowe sposoby prowadzenia badań.

Badacze wykorzystujący takie portale naukowe mogą traktować je jako dedykowane środowiska pracy, gdzie organizują swoje dane i przeprowadzają na nich skomplikowane obliczenia za pomocą oprogramowania danej dziedziny nauki, co ważne, w sposób specyficzny dla tej dziedziny.

Portale bazujące na technologii InSilicoLab są zaprojektowane tak, aby gromadzić w jednym miejscu całość narzędzi, których badacze potrzebują dla obliczeń *in silico*. To oznacza:

- możliwość organizowania danych będących przedmiotem badań lub ich wynikiem – w tym:
 - ułatwienie procesu przygotowania danych wejściowych dla obliczeń,
 - możliwość opisu i kategoryzacji danych wejściowych i wyjściowych za pomocą metadanych umożliwiających później szybkie odszukanie właściwych danych,
 - wyszukiwanie i przeglądanie dostępnych danych i opisujących je metadanych,
- łatwe uruchamianie nawet skomplikowanych, długich i wymagających wielu obliczeń, eksperymentów użytkownika.



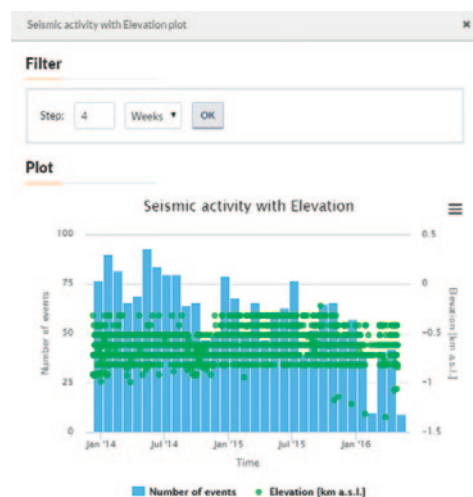
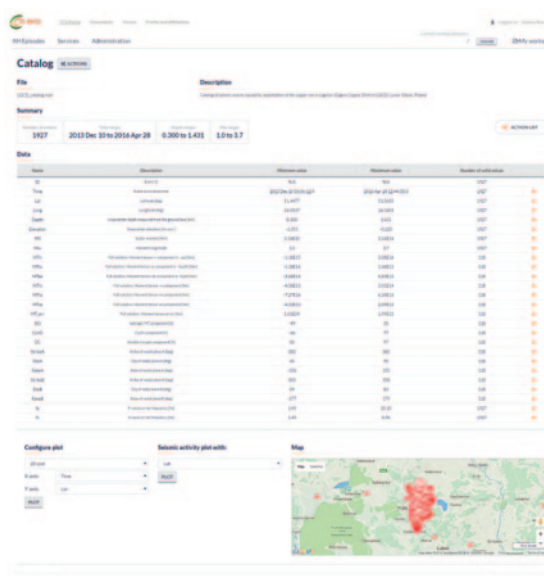
Cele platformy

To co wyróżnia technologię InSilicoLab, to dążenie do jak największej użyteczności budowanych za pomocą środowiska narzędzi, a więc jak najdokładniejszego dopasowania do potrzeb danej dziedziny nauki. Oznacza to przyjazność portalu naukowego dla użytkownika, ale co ważniejsze, użyteczność dla rozwiązywania rzeczywistych problemów naukowych. Tak postawione zadanie wymaga przygotowania dedykowanych narzędzi dziedzinowych umożliwiających rozwiązywanie specyficznych problemów interesujących badaczy, a zatem nie „ogólnego” środowiska umożliwiającego rozwiązanie każdego problemu badawczego z danej dziedziny. Zbudowanie takiego uniwersalnego i wszechstronnego narzędzia badawczego nie jest możliwe bez poświęcenia wygody i intuicyjności użytkowania, a brak tej ostatniej stanowi olbrzymią barierę w pozyskiwaniu nowych użytkowników. Przygotowanie dedykowanego portalu naukowego musi siłą rzeczy bazować na specyficznych problemach badawczych w ramach danej dziedziny. Cały portal jest natomiast sumą poszczególnych części składowych: dziedzinowych narzędzi, oprogramowania i niezbędnych serwisów, a także bibliotek należących do samej technologii InSilicoLab, stanowiących wspólną bazę dla wielu dziedzinowych portali.

Aplikacje dziedzinowe

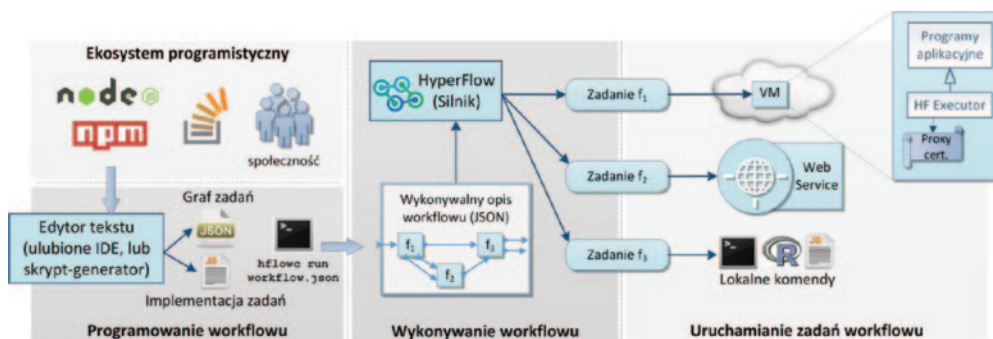
Każdy portal naukowy bazujący na technologii InSilicoLab jest dopasowany do specyficznych wymagań danej dziedziny lub też tylko do pewnej klasy jej problemów. Narzędzia wspólne środowiska dostarczają usług do zarządzania danymi użytkowników, ich kategoryzacji, opisu metadanymi i mechanizmów pozwalających określić źródło danych. Od strony infrastruktury zaś umożliwiają prowadzenie na niej obliczeń. Na tych narzędziach zbudowane są dziedzinowe modele danych i metadanych, scenariusze badań i specyficzne interfejsy użytkownika.

<http://insilicolab.cyfronet.pl>





HyperFlow jest narzędziem, które umożliwia kompozycję aplikacji naukowych w złożone przepływy, tzw. *workflows naukowe*. HyperFlow wspomaga tworzenie przepływów, rozmieszczanie ich w chmurze obliczeniowej oraz ich uruchamianie.



Programowanie workflowów naukowych

W HyperFlow workflow jest opisywany jako graf zadań (nazywanych *procesami*) w prostym formacie opartym o JSON. Zadania workflowu reprezentują procedury naukowe – kolejne kroki w przepływie naukowym. Zadania te mogą być albo zaimplementowane w języku JavaScript, albo odwzorowane na programy wykonywalne. W pierwszym przypadku zadania workflowu są wykonywane przez `Node.js`, szeroko wykorzystywane środowisko uruchomieniowe dla języka JavaScript. Dzięki temu programowanie workflowu osadzone jest w dojrzałym ekosystemie programistycznym (społeczność, narzędzia, biblioteki, zasoby sieciowe), nie zaś w środowisku specyficznym dla danego systemu, mało znanym i wspieranym w ograniczonym stopniu. Dzięki możliwości niskopoziomowego programowania zadań workflowu, mogą one z łatwością wywoływać zewnętrzne Web Service’y, czy też wykonywać lokalne komendy, w ramach jednego przepływu.

W drugim przypadku twórca workflowu nie musi implementować kodu JavaScript, a jedynie przypisać każde zadanie workflowu do uprzednio utworzonej maszyny wirtualnej, na której zainstalowane są aplikacje naukowe, oraz określić komendę, która zostanie wykonana w przypadku uruchomienia danego zadania.

Dzięki wsparciu tych dwóch podejść do programowania workflowu, HyperFlow jest przydatny zarówno doświadczonym programistom / inżynierom, którzy pragną mieć możliwość programowania niskopoziomowego, jak i naukowcom dziedzinowym, którzy chcą jedynie komponować przepływy z uprzednio istniejących modułów.

Rozmieszczanie w chmurze

HyperFlow automatycznie rozmieszcza workflow w chmurze obliczeniowej. Aby to było możliwe, użytkownik musi jedynie przygotować plik konfiguracyjny, w którym określone jest odwzorowanie zadań workflowu na dostępne obrazy maszyn wirtualnych. Następnie użytkownik wykonuje prostą komendę *hflowc setup*, w wyniku której tworzone są instancje maszyn wirtualnych w chmurze. Instancje te zawierają komponenty środowiska uruchomieniowego HyperFlow oraz aplikacje naukowe uruchamiane z zadań workflowu.

Wykonanie workflowu

Po utworzeniu instancji workflowu w chmurze, użytkownik może ją uruchomić przy pomocy komendy *hflow run <workflow_directory>*. Każdy workflow uruchomiony jest ze swoją własną instancją środowiska uruchomieniowego HyperFlow. Powoduje to, że poszczególne workflowy są od siebie izolowane, co zwiększa bezpieczeństwo i niezawodność.

Środowisko uruchomieniowe HyperFlow (tzw. *HyperFlow Executor*) automatycznie przesyła dane wejściowe z katalogu użytkownika na instancje maszyn wirtualnych, wykonuje aplikacje naukowe, a następnie przesyła dane wyjściowe z powrotem do katalogu użytkownika. HyperFlow obsługuje kilka mechanizmów transferu danych, m.in. sieciowy system plików, gridftp oraz Amazon S3.

Aplikacje

HyperFlow jest częścią kilku większych systemów, w których jest wykorzystywany do implementacji szeregu aplikacji. W infrastrukturze PLGrid HyperFlow jest zintegrowany jako system zarządzania workflowami naukowymi, który umożliwia kompozycję naukowych workflowów i ich uruchamianie w chmurze. Przykładową aplikacją wspieraną przez HyperFlow jest solwer dla problemów opartych na metodzie elementów skończonych, który może być użyty do rozwiązywania wielu różnorodnych problemów naukowych. W projekcie Paasage (<http://www.paasage.eu>) HyperFlow jest integrowany jako silnik wykonawczy dla aplikacji naukowych uruchamianych w kontekście środowiska *multi-cloud*. W projekcie ISMOP (<http://www.ismop.edu.pl>) HyperFlow jest częścią systemu podejmowania decyzji w sytuacjach zagrożenia powodziowego. Hyperflow będzie miał również zastosowanie w obliczeniach parametrycznych i przetwarzaniu typu workflow, jako zamiennik technologii Scalarm.

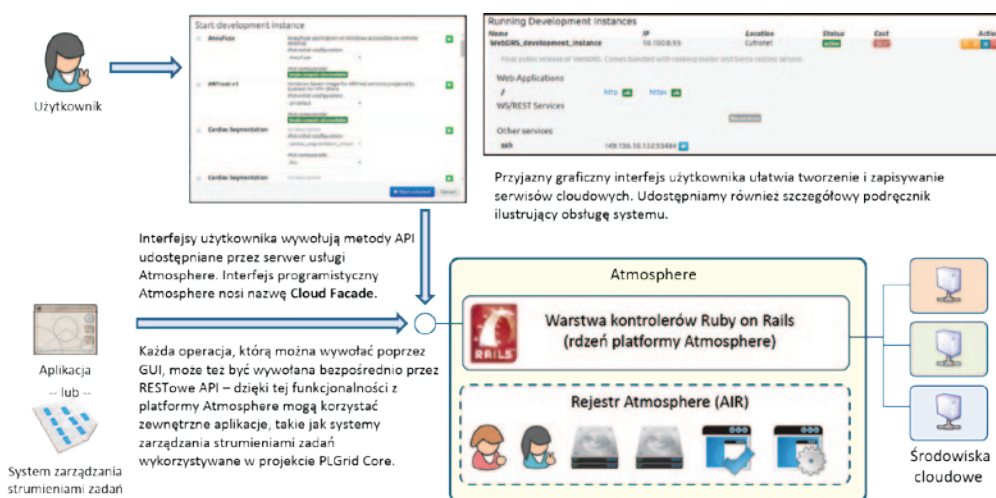
Kontakt

HyperFlow jest rozwijany i utrzymywany przez zespół DICE (<http://dice.cyfronet.pl>). Zapraszamy do zgłaszania wszelkich pytań oraz sugestii.

Kod źródłowy HyperFlow oraz podręczniki instalacji i użytkowania są dostępne na stronie <https://github.com/hyperflow-wms/hyperflow>.

Atmosphere

Platforma cloudowa Atmosphere jest zintegrowanym systemem umożliwiającym zarządzanie oraz interakcję z zasobami cloudowymi udostępnianymi w infrastrukturze PLGrid. Atmosphere to hybrydowe środowisko użytkownika umożliwiające integrowanie zasobów cloudowych dostarczanych przez rozmaite instytucje (w tym instytucje komercyjne). Platforma prezentuje swoim użytkownikom jednolitą, spójną przestrzeń zasobów, na których użytkownicy PLGrid mogą instalować i uruchamiać aplikacje naukowe wykorzystujące zasoby obliczeniowe dużej mocy. Oprócz narzędzi dla użytkowników aplikacji, udostępniamy również bogaty wybór narzędzi dla programistów oraz administratorów infrastruktur obliczeniowych.



Abstrakcja usług cloudowych

Podstawowym celem platformy Atmosphere jest sprawienie, by interakcja z zasobami cloudowymi była prosta i przejrzysta, zarówno dla początkujących jak i zaawansowanych użytkowników. Atmosphere jest w stanie instalować maszyny wirtualne w cloudzie, tworzyć obrazy maszyn oraz wspomagać współdzielenie usług obliczeniowych w zespołach PLGrid bez konieczności korzystania z niskopoziomowych bibliotek i interfejsów programistycznych. Dla wszystkich trzech klas użytkowników – administratorów, twórców aplikacji oraz użytkowników aplikacji – udostępniane są wygodne graficzne interfejsy użytkownika, a także niskopoziomowe interfejsy programistyczne umożliwiające integrowanie platformy Atmosphere z zewnętrznymi usługami, aplikacjami naukowymi oraz platformami zarządzania strumieniami zadań (systemami typu *workflow*).

Bezpieczeństwo

Platforma Atmosphere jest w pełni zintegrowana z mechanizmami uwierzytelniania oraz autoryzacji dostępu wbudowanymi w infrastrukturę PLGrid. Każdy z użytkowników PLGrid może zawnioskować o dostęp do zasobów cloudowych zapisując się do zespołu **plgg-cloud2**. Po akceptacji wniosku użytkownicy mogą tworzyć, wykorzystywać i współdzielić maszyny wirtualne w kontekście swoich własnych zespołów badawczych, a także na potrzeby indywidualnych projektów naukowych. Widoczność maszyn wirtualnych (oraz skojarzonych z nimi obrazów) jest automatycznie ograniczana do zespołu PLGrid, w ramach którego dane maszyny zostały zainicjalizowane. Atmosphere dba o automatyczne monitorowanie wykorzystania zasobów sprzętowych przez wszystkie maszyny wirtualne.

Zasoby

Atmosphere jest w stanie integrować cloudowe zasoby obliczeniowe udostępniane w różnych technologiach. W ramach infrastruktury PLGrid w ACK Cyfronet AGH udostępniono zasoby obliczeniowe zarządzane przez popularny stos cloudowy OpenStack. Atmosphere potrafi również pozyskiwać zasoby od publicznych dostawców infrastruktur cloudowych, takich jak Amazon, RackSpace, Google Compute, itp. Opisana funkcjonalność nie wymaga od użytkowników opanowania dedykowanych bibliotek programistycznych ani interfejsów API.

Aplikacje

Oprócz zestawu szablonów systemów operacyjnych, platforma cloudowa PLGrid obsługuje szereg aplikacji naukowych. Poprzednia wersja platformy Atmosphere, wdrożona w projekcie VPH-Share, była z powodzeniem wykorzystywana przez około 25 niezależnych zespołów badawczych dla symulacji z zakresu Virtual Physiological Human w ramach konsorcjów VPH-Share oraz VPH-DARE, a także przez szereg samodzielnych zespołów współpracujących z ACK Cyfronet AGH oraz członków projektu EurValve, którego uczestnikiem był również Cyfronet.

<https://cloud.plgrid.pl>

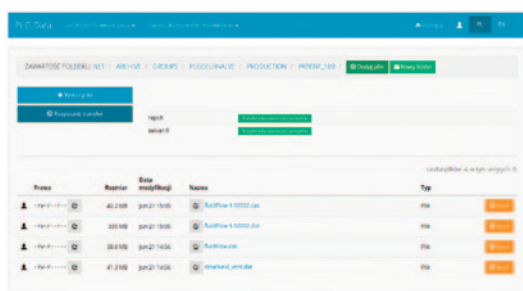


Proste narzędzie zarządzania plikami na klastrze obliczeniowym

Usługa PLG-Data pozwala, za pośrednictwem przeglądarki WWW, zarządzać swoimi zbiorami danych, zgromadzonymi na zasobach dyskowych infrastruktury PLGrid. Dzięki niej można oglądać, pobierać, dodawać i usuwać pliki oraz foldery. Ułatwia także zarządzanie prawami dostępu dla innych osób, członków zespołu badawczego lub współpracowników spoza tej grupy. Posiada wsparcie dla klastrów Prometheus i Zeus.

Zestaw funkcjonalności, na które pozwala narzędzie, to m.in.:

- pobieranie plików z klastra na dysk lokalny,
- dodawanie nowych i usuwanie istniejących zbiorów,
- zmiana nazw plików i folderów, jak również zmiana uprawnień dostępu do nich,
- szybka nawigacja pomiędzy katalogami: domowym, roboczym i zespołowym, dzięki poręcznemu menu skrótów,
- sprawny podgląd plików graficznych bez konieczności pobierania ich na dysk.



Dzięki odpowiedniej konstrukcji adresu URL do poszczególnych zbiorów, narzędzie pozwala łatwo przekazać ich lokalizację innej osobie, np. poprzez wklejenie wprost treści URL z paska adresu przeglądarki do wiadomości e-mail czy komunikatora. Adresat znajdzie współdzielony katalog lub pobierze plik jednym kliknięciem – pod warunkiem, że posiada odpowiednie uprawnienia dostępu.

Usługa jest zabezpieczona szyfrowanym protokołem HTTPS (na odcinku połączenia komputera użytkownika z portalem PLG-Data) oraz protokołem GridFTP – pomiędzy portalem usługi oraz klastrami obliczeniowymi. Zastosowanie tych technik pozwala użytkownikom korzystać z niej w sposób bezpieczny oraz, z punktu widzenia innych współużytkowników PLGrid, ogranicza ich możliwości do zakresu przyznanego przez właścicieli danych. Innymi słowy, osoba korzystająca z usługi PLG-Data nie nabywa żadnych dodatkowych uprawnień w dostępie do zbiorów danych poza tymi, które już posiada.

Logowanie do narzędzia odbywa się poprzez parę login-hasło PLGrid lub certyfikat p12 załadowany w przeglądarce. Usługa jest dostępna w polskiej i angielskiej wersji językowej, posiada także rozbudowany interfejs programistyczny (API), który pozwala na jej użycie przy budowie platform, narzędzi czy innych usług wykorzystujących zasoby infrastruktury PLGrid.

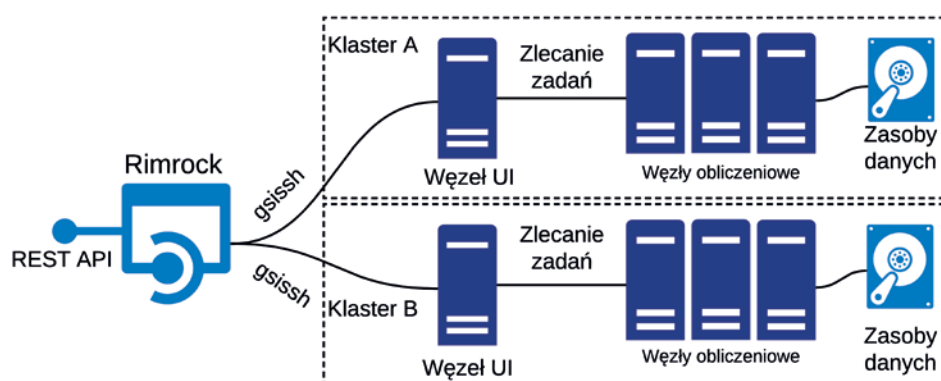
Adres narzędzia PLG-Data: <https://data.plgrid.pl>



Rimrock to usługa infrastruktury PLGrid umożliwiająca kontrolę obliczeń naukowych i zarządzanie otrzymywanymi za ich pomocą rezultatami dzięki nowoczesnym interfejsom dostępowym opartym o REST. REST (ang. *Representational State Transfer*) jest swego rodzaju wzorcem projektowania architektur rozproszonych aplikacji. Dzięki wykorzystaniu REST możliwy jest łatwy dostęp do usług, aplikacji i zaawansowanych skryptów.

Dostępność dla różnorodnych zastosowań

Zastosowanie interfejsów REST przy budowie usługi rimrock umożliwia wykorzystanie funkcjonalności usługi niezależnie od języka programowania, który został wybrany do tworzenia systemów wykorzystujących zasoby obliczeniowe. Można więc w oparciu o rimrock budować aplikacje internetowe i desktopowe, jak również tworzyć zaawansowane skrypty obliczeniowe (np. za pomocą języka powłoki *Bash* i polecenia *curl*). Ciekawym rozwiązaniem oferowanym przez usługę jest możliwość tworzenia aplikacji internetowych uruchamianych w całości w przeglądarce użytkownika, przy zminimalizowaniu udziału strony serwerowej.



Integracja z wieloma systemami zarządzania obliczeniami

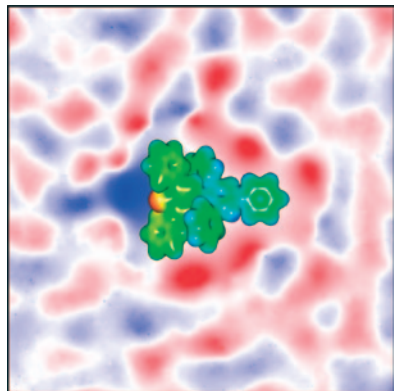
Usługa rimrock wykorzystuje kilka systemów zarządzania zadaniami (jak PBS/TORQUE lub Slurm), dzięki czemu zapewnia wsparcie dla ich unikatowych własności. Pozwala to na łatwą integrację aplikacji już istniejących w rozwijanych systemach. Dostęp do rezultatów obliczeń jest ułatwiony poprzez ukrycie wewnętrznego protokołu przesyłania plików (*GridFTP*) i grupowanie wyników zadań.

Bezpieczeństwo danych

Wymiana danych z usługą jest realizowana za pomocą bezpiecznego połączenia HTTPS, a do autoryzacji użytkowników wykorzystywany jest tymczasowy certyfikat użytkownika (tzw. *proxy*).

<https://submit.plgrid.pl>

Chemia i Biologia – struktura elektronowa i dynamika molekularna



A. Eilmes, P. Kubisiak: Potencjał elektrostatyczny cieczy jonowej w otoczeniu rozpuszczonej cząsteczki barwnika

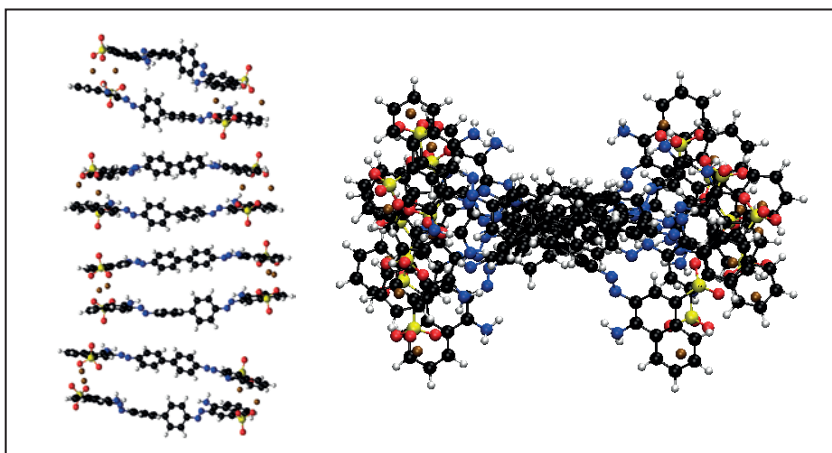
Współczesna chemia obliczeniowa wymaga ciągłego zwiększania zasobów, które umożliwiają przeprowadzenie symulacji dla coraz większych układów (ważnych m.in. w nanotechnologii i naukach biologicznych) oraz poprawę dokładności uzyskanych wyników. Na szczęście, ciągły rozwój technologii w połączeniu z profesjonalnym oprogramowaniem oferowanym przez Cyfronet wychodzi naprzeciw rosnącym wymaganiom i umożliwia przeprowadzanie różnorodnych obliczeń z zakresu chemii kwantowej.

Superkomputery *Zeus* oraz *Prometheus* składają się z węzłów obliczeniowych dysponujących nawet 1,5 TB RAM i 64 rdzeniami, co umożliwia wykonywanie obliczeń kwantowochemicznych wymagających ogromnych zasobów pamięci lub dużej liczby rdzeni ze współdzieloną pamięcią operacyjną. Co więcej, wszystkie węzły superkomputera połączone są szybkim interfejsem InfiniBand o wysokiej przepustowości i krótkim czasie dostępu, co przekłada się na zwiększenie szybkości wykonywania obliczeń, jeśli są one rozproszone pomiędzy węzły obliczeniowe. Wiele programów do symulacji chemicznych wymaga również szybkiego dostępu do systemów dyskowych, który jest możliwy dzięki wykorzystaniu rozproszonego systemu plików Lustre.

Wydajność zadań z chemii obliczeniowej zależy także od odpowiedniej konfiguracji oprogramowania naukowego oraz umiejętności jego użycia. Nasz zespół administratorów posiada niezbędną wiedzę, umiejętności oraz duże doświadczenie w instalacji oraz efektywnym wykorzystaniu dostępnych aplikacji i narzędzi. W naszym szerokim portfolio można odnaleźć wiele programów umożliwiających różnorodne obliczenia chemiczne. Wśród nich:

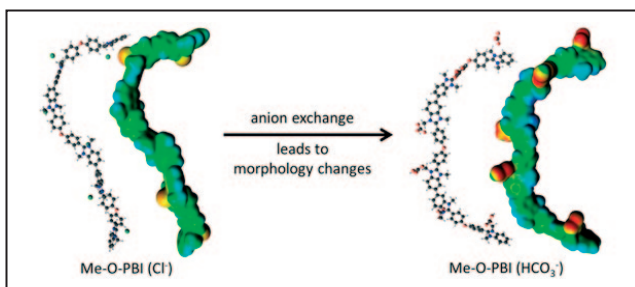
- Wszechstronne i powszechnie używane programy kwantowochemiczne takie jak **Gaussian**, **GAMESS US**, **NWChem**, **Schrödinger**, **Q-Chem** czy **TURBOMOLE**, dzięki którym można przeprowadzić obliczenia struktury elektronowej oraz wielu właściwości różnorodnych systemów molekularnych wykorzystując metody *ab initio*, pół-empiryczne oraz teorie funkcjonałów gęstości.
- Pakiety **Molpro**, **CFOUR** i **Dalton** do analizy systemów molekularnych z dużą dokładnością, w oparciu o zaawansowane schematy obliczeniowe, takie jak metody sprzężonych klastrow (CCSD(T)) oraz multireferencyjne (MCSCF).
- Pakiet Amsterdam Modeling Suite (**ADF**, **DFTB**, **MOPAC**, **COSMO-RS**) udostępniający metody do analizy różnorodnych właściwości systemów molekularnych (w szczególności spektroskopowych, takich jak widma NMR czy ESR) poprzez m.in. wykorzystanie relatywistycznego przybliżenia ZORA oraz metody COSMO-RS. Dodatkowo, rozbudowane GUI aplikacji **ADF** (**ADFInput**, **ADFView**, etc.) ułatwia pracę naszym użytkownikom.

- Znacząca część dostępnych aplikacji może być używana do obliczeń właściwości systemów periodycznych. Wśród nich dostępne są **BAND**, **Quantum ESPRESSO** oraz **SIESTA**.
- **Desmond**, **Gromacs**, **Amber**, **LAMMPS**, **NAMD**, **Tinker-HP**, **CPMD**, **CP2K** i **Terachem** przeznaczone do symulacji wykorzystujących mechanikę molekularną oraz dynamikę molekularną dla systemów zbudowanych nawet z setek tysięcy atomów.



O. Klimas: Obliczona optymalna struktura klastra ośmiu cząsteczek czerwieni Kongo widziana z różnych perspektyw

Wykorzystanie kart graficznych do przeprowadzania obliczeń (**GPGPU**) w wielu dziedzinach nauki pozwala na wielokrotne przyspieszenie wykonywanych zadań. W naszym Centrum część węzłów umożliwia takie obliczenia poprzez użycie jednocześnie nawet ośmiu kart GPGPU wykorzystujących środowisko programistyczne **CUDA**. Wiele pakietów dostępnych w Cyfronecie umożliwia obliczenia GPGPU, wśród nich oprogramowanie chemiczne: **GAMESS**, **Terachem**, **NAMD**, **Quantum ESPRESSO**. Ekspert z ACK Cyfronet AGH współpracują z wieloma grupami tworzącymi i rozwijającymi powyższe pakiety. Warto dodać, że pracownicy Cyfronetu włożyli wiele wysiłku, aby dopracować środowiska obliczeniowe dla naszych użytkowników.



Potencjał elektrostatyczny cząsteczki tworzącej membranę jonowymienną. Opublikowane przez W. Germer, J. Leppin, C. Kirchner, H. Cho, H. Kim, D. Henkensmeier, K. Lee, M. Brela, A. Michalak oraz A. Dyck w Macromol. Mater. Eng. 2015, 300, 497–509

Uczenie maszynowe (ML) i sztuczna inteligencja (AI)

Analiza danych przyspieszona przez sztuczną inteligencję czyni wielkie postępy w wielu dziedzinach badań, w tym w naukach o materiałach, naukach przyrodniczych, językoznawstwie i naukach społecznych. Zdolność sieci neuronowych do uczenia się na podstawie złożonych danych może znacznie poprawić analizę danych, klasyfikację i wykrywanie wzorców, z potencjalnymi zastosowaniami w wielu systemach, m.in. w rozpoznawaniu obrazów, przetwarzaniu języka i optymalizacji.

Centrum superkomputerowe Cyfronet wychodzi naprzeciw tym wyzwaniom i przygotowało kilka pakietów:

PYTORCH

PyTorch to pakiet, specjalna biblioteka dla języka Python, oparta na bibliotece Torch. Pozwala na implementację złożonych algorytmów Uczenia Głębokiego z obszaru Przetwarzania Języka Naturalnego, przetwarzania wideo i obrazów oraz wielu innych. Nadaje się bardzo dobrze do modelowania nowych architektur z dziedziny uczenia maszynowego z nastawieniem na eksperymenty.

TensorFlow pozwala podobnie jak Pytorch na implementację modeli bazujących na paradygmacie przepływu tensorów. Ze względu na swój charakter oraz statyczny graf reprezentacji umożliwia wydajną optymalizację treningu oraz inferencji modeli z uwzględnieniem platformy obliczeniowej.



Keras

Keras to biblioteka do projektowania modeli neuronalnych. Stanowi zewnętrzne API dla silników bazujących na TensorFlow, Microsoft Cognitive Toolkit, Theano lub PlaidML. Keras został zaprojektowany tak, aby był przyjazny dla użytkownika – jest aplikacją modułową i rozszerzalną.

Scikit-learn to biblioteka dla języka Python. Posiada różne algorytmy klasyfikacji, regresji i grupowania, w tym obsługę maszyn wektorowych, losowych lasów, zwiększania gradientu, k-średnich i DBSCAN. Został zaprojektowany do współdziałania z bibliotekami numerycznymi i naukowymi Pythona: NumPy i SciPy.



SchNet to architektura głębokiego uczenia (deep learning), która umożliwia modelowanie molekularne materiałów na poziomie atomowym.

Horovod to zestaw narzędzi dla TensorFlow, Keras, PyTorch i MXNet. Głównym założeniem Horovod jest uczynienie procesu głębokiego uczenia (deep learning) rozproszonym, aby zwiększyć wydajność obliczeń i równocześnie ułatwić korzystanie z niego.

Wizualizacja danych, POV-Ray/ScPovPlot3D

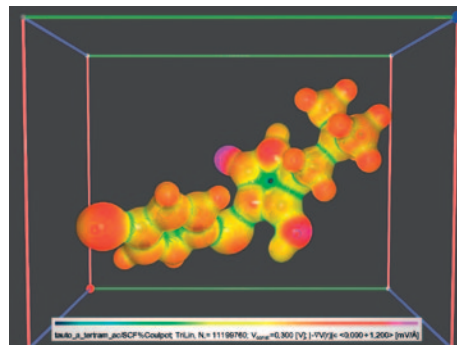
Wizualizacja danych umożliwia analizę i zrozumienie wyników nawet bardzo złożonych obliczeń, szczególnie, jeśli ich forma finalna jest wielowymiarowa bądź uwzględnia czynnik czasu. Większość aplikacji do obliczeń numerycznych posiada jakiś moduł generujący ich wizualizację. W Python'ie jest dostępny moduł matplotlib czy VTK+, biblioteki graficzne posiadają także Matlab czy R. Podobnie wygląda sytuacja w przypadku programów do geowizualizacji (GIS) czy programów do obliczeń chemicznych. Niestety, niezależnie od tego, jak bardzo programy te są dopracowane, to wynik ich działania jest ograniczony iloczynem kartezjańskim dostępnych (a przy tym zgodnych) opcji.

Usunięcie tego ograniczenia – choćby tylko w celu stworzenia prototypu stylu wizualizacji do późniejszego wdrożenia w dedykowanym pakiecie – jest możliwe, wymaga jednak skorzystania z programu graficznego ogólnego przeznaczenia. Dobrymi przykładami są tu 3DMax, Blender czy POV-Ray. Jednak tylko ten ostatni wyposażony jest w domenowy język opisu sceny (SDL) o charakterze skryptowym, który pozwala na programowe, nieinteraktywne tworzenie wizualizacji. Posługiwanie się niezliczoną liczbą opcji języka SDL wymaga dość długiej nauki, dlatego napisano dedykowane API, w formie zestawu specjalizowanych modułów – pakiet ScPovPlot3D. Nie jest to projekt zamknięty, wciąż dopisywane są kolejne rozszerzenia, można zatem mówić o dojrzałej, działającej wersji beta. Aktualnie projekt znajduje się w wersji 4.0 i jest hostowany na GitHub'ie (URL: <https://github.com/JustJanush/Plot3Dv4>) – wieloplatformowe API wymaga POV-Ray'a w wersji min. 3.7.

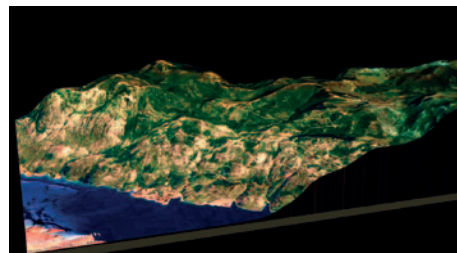
Najważniejsze moduły to:

- VectorField.inc – hybrydowa wizualizacja pól wektorowych z użyciem widżetów i/lub rurek strumienia pola,
- Potential.inc – hybrydowa wizualizacja pól skalarnych, na siatkach regularnych i nieregularnych z interpolacją trilinearną albo sześcienną Catmull-Rom'a,
- BPatchSurf.inc – hybrydowa wizualizacja powierzchni w oparciu o dane na siatkach regularnych i nieregularnych z zaimplementowanym prostym krigingiem (KDE),
- Mesh2Surf.inc – hybrydowa wizualizacja danych określonych na regularnych siatkach 2D ($z = f(x, y)$),
- TextExt.inc – rozszerzone formatowanie tekstu 3D, zorientowane na prezentację równań.

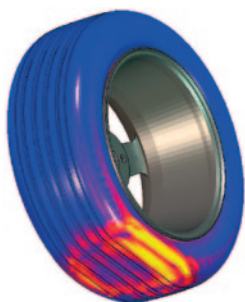
W razie potrzeby developer pakietu służy wsparciem technicznym. Informacje kontaktowe: <https://skos.agh.edu.pl/osoba/janusz-opila-2390.html>.



Janusz Opila: konfiguracja pola elektrostatycznego wokół molekuly polimeru. Ukazano powierzchnię stałego potencjału z aproksymacją trilinearną, kolor koduje moduł natężenia pola elektrostatycznego



Janusz Opila: Wizualizacja terenu uzyskana w oparciu o dane wysokościowe zebrane na nieregularnej siatce i tekstury pozyskane z aplikacji GoogleEarth-Pro (okolice Karlobagu, Chorwacja). Opracowanie własne, DOI: 10.23919/MIPRO.2018.8400037



Aplikacje typu CAD (projektowanie wspomagane komputerowo) oraz CAE (komputerowe wspomaganie prac inżynierskich) stanowią dziś nieodłączną część procesu projektowania i budowy niemal wszystkiego – od części samochodowych po budynki. Dzięki symulacjom komputerowym inżynierowie mogą dokonywać oceny wytrzymałościowej elementów maszyn i konstrukcji, wykonywać liniowe i nieliniowe analizy strukturalne z uwzględnieniem odkształceń, zjawisk kontaktowych, plastyczności, hipersprężystości, itp. Istotne miejsce w dziedzinie nauk ścisłych zajmują również symulacje dotyczące: spalania, turbulencji, przepływów wielofazowych, reakcji chemicznych, radiacji i mechaniki płynów.

Wszystkie powyższe zagadnienia mogą być badane przez użytkowników dzięki specjalistycznemu oprogramowaniu, np. ANSYS, ABAQUS, FLUENT, MARC czy OPERA.

ANSYS jest rozbudowanym środowiskiem, posiadającym intuicyjny interfejs graficzny, wspierającym naukowców z prawie wszystkich nauk ścisłych oraz przemysłu. Uzyskiwane wyniki charakteryzują się dużą dokładnością. Rezultaty prezentowane są w postaci tekstowej i graficznej, np. diagramy, izopowierzchnie, deformacje. Możliwości obliczeniowe ANSYS-a są bardzo duże, obejmują one m.in. analizę harmoniczną i spektralną, lepkosprężystość, drgania własne, statykę i dynamikę.

ABAQUS znajduje szerokie zastosowanie w przemyśle, a wykonywanie obliczeń opiera się o metodę elementów skończonych. Użytkownik opracowuje kombinacje elementów skończonych, materiałów, procedur analizy i sekwencji obciążeń, zgodnie ze swoimi wymaganiami.

FLUENT umożliwia badanie i symulację zachowania cieczy, co znajduje zastosowanie przy projektowaniu urządzeń wykorzystujących lub transportujących płyny. Może być wykorzystywany w dziedzinie chemii, elektroniki, metalurgii, energetyki cieplnej, biomedycyny i innych.

MARC jest środowiskiem ogólnego zastosowania przeznaczonym do wykonywania obliczeń nieliniowej analizy elementów skończonych w celu wykonania symulacji zachowania produktu przy zmiennych warunkach statycznych, dynamicznych oraz z uwzględnieniem wszystkich rodzajów nieliniowości, w tym geometrycznej, materiałowej oraz warunków brzegowych i kontaktu. Jest to rozwiązanie posiadające możliwość symulacji produkcji i testowania wytworzonego produktu w celu wykrycia zniszczeń, uszkodzeń, awarii oraz propagacji pęknięć. Wszystko to można połączyć z możliwą analizą termiczną, elektryczną, magnetyczną i strukturalną.

OPERA jest oprogramowaniem używającym metody elementów skończonych do projektowania i optymalizacji urządzeń elektromagnetycznych, zarówno w dwóch jak i trzech wymiarach. Pozwala na wyszukiwanie rozwiązań numerycznych dla problemów z wielu dziedzin nauki: elektrostatyki, magnetostatyki, elektromagnetyzmu. Oprogramowanie umożliwia projektowanie różnych urządzeń elektrycznych: transformatorów, silników, przełączników, skanerów MRI i lamp rentgenowskich. To potężne narzędzie do wirtualnego prototypowania, dzięki któremu proces projektowania trwa o wiele krócej.

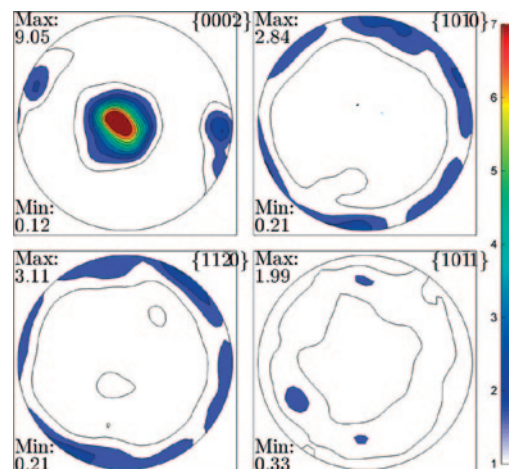
Aplikacje matematyczne

Dzięki aplikacjom matematycznym nawet bardzo rozbudowane i skomplikowane obliczenia mogą zostać wykonane w krótkim czasie. Użytkownicy ACK Cyfronet AGH mają dostęp do programów wspomagających wykonywanie zaawansowanych zadań z algebry, analizy, kombinatoryki, statystyki, teorii liczb, geometrii jak i innych działów matematyki. Wykonywanie działań takich jak całkowanie, różniczkowanie, obliczenia symboliczne, aproksymacja i interpolacja, Transformata Fouriera i Laplace'a, operacje macierzowe, przetwarzanie sygnałów cyfrowych, itp. staje się o wiele łatwiejsze. Przy pomocy odpowiednich narzędzi możliwa jest nawet wizualizacja wyników działań na zbiorach liczb rzeczywistych i/lub zespolonych. Niektóre programy pozwalają dodatkowo na tworzenie wykresów interaktywnych 2D bądź 3D. Wszystkie te funkcjonalności są bardzo istotne, ponieważ w pracy naukowców kluczowe jest często przygotowanie precyzyjnego modelu, jak najdokładniej opisującego analizowane zagadnienia.

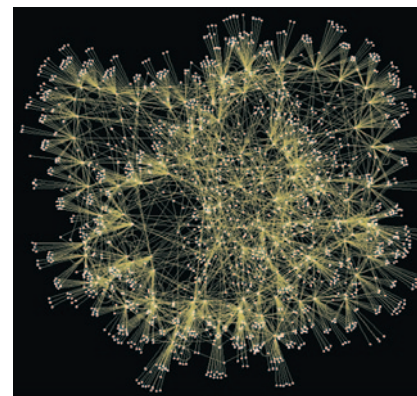
Przykładem środowiska znajdującego zastosowanie w wyżej opisanych zagadnieniach jest **MATLAB**. Dzięki swoim dodatkowym modułom pozwala na wykonywanie obliczeń z zakresu modelowania finansowego, równań różniczkowych cząstkowych (PDE), czy optymalizacji wielokryterialnej, liniowej oraz nieliniowej. Inne zastosowania obejmują m.in. analizę danych geograficznych i generowanie na ich podstawie map, modelowanie biologicznych systemów dynamicznych i systemów hydraulicznych, czy analizę i wizualizację danych genomicznych oraz proteomicznych. Możliwe jest również użycie środowiska symulacyjnego Simulink, służącego do budowania modeli symulacyjnych z bloków, wizualizacji i analizy danych, bez potrzeby tradycyjnego programowania.

Obok tego rozbudowanego środowiska, użytkownicy znajdą w oprogramowaniu oferowanym przez Centrum także użyteczną aplikację, **MATHEMATICA**, dedykowaną do prowadzenia obliczeń równoległych ze zdefiniowaną dokładnością, do obliczeń symbolicznych i numerycznych. Zaletą tego pakietu jest także posiadanie narzędzia do automatycznego wykrywania i poprawiania błędów oraz możliwość zlecania zadań na kilku komputerach jednocześnie.

Innym uniwersalnym i interakcyjnym pakietem matematycznym jest **MAPLE**. Stosuje się go m.in. do upraszczania wyrażeń i dokładnej reprezentacji wielkości matematycznych. Oferuje możliwość korzystania z baz danych, generowania kodu innych języków programowania, tworzenia pokazów slajdów zawierających wykonywane zadania wraz z własnymi komentarzami, czy komunikację z aplikacjami MATLAB i systemami CAD.

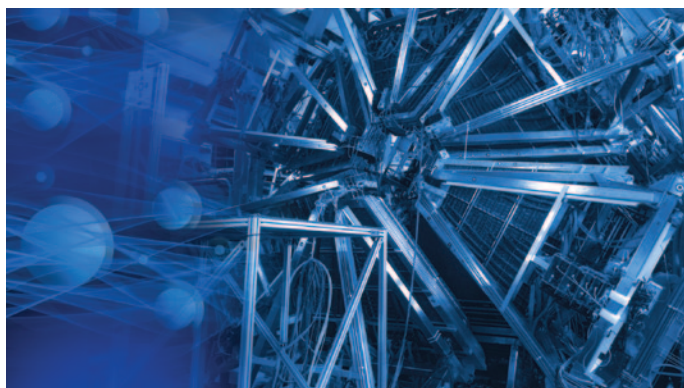


Bartosz Sułkowski:
Wyniki symulacji
tekstury Zn
po wytlaczaniu
hydrostatycznym
w 250 °C



Rafał Rak: Sieć połączeń
jednominutowych stóp
zwrotu dla spółki KGHM
(polska spółka giełdowa)

Intensywnie rozwijające się badania naukowe wymagają coraz bardziej zaawansowanych narzędzi. Wśród nich ogromną rolę pełnią narzędzia informatyczne, wspomagające prowadzenie efektywnych badań od momentu ich projektowania, aż po opracowywanie wyników. Cyfronet, poprzez śledzenie najnowszych rozwiązań i tworzenie własnych opracowań, stara się wypełniać ważny obszar swojej misji wsparcia nauki. Na te potrzeby powstały specjalistyczne laboratoria.



Laboratorium Obliczeń Kwantowych

Laboratorium powstało w celu prowadzenia prac badawczych w zakresie wykorzystania komputerów kwantowych w obliczeniach i wsparcia obliczeń klasycznych akceleratorami kwantowymi.

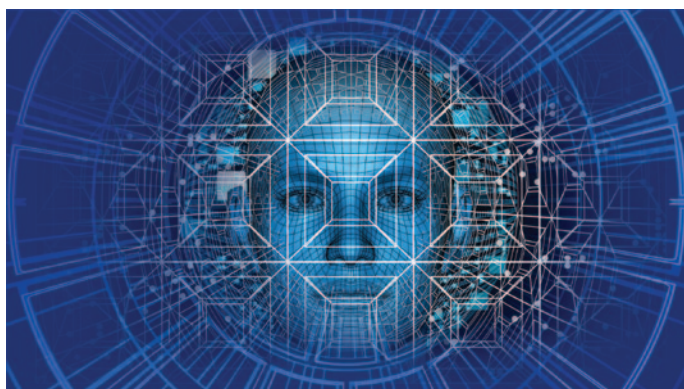
Jednym z kluczowych zadań jest śledzenie rozwoju technologii obliczeń kwantowych oraz dostępnych platform akceleratorów kwantowych celem ich wyko-

rzystania w dedykowanych usługach ACK Cyfronet AGH. Współpracujemy z innymi podmiotami badawczymi i partnerami przemysłowymi, zarówno jako konsorcjanci wspólnych inicjatyw, jak i realizując prace zlecone.

W oparciu o własne kompetencje oraz wymianę wiedzy eksperckiej z siecią partnerów, nasz zespół pracuje nad rozwiązaniem problemów stojących na drodze do szerszego i bardziej efektywnego wykorzystania akceleratorów kwantowych w obliczeniach na rzecz nauki i gospodarki.

Działamy również na rzecz popularyzacji obliczeń wykorzystujących akceleratorów kwantowych oraz merytorycznego wsparcia użytkowników. W tym zakresie przygotowujemy niezbędną dokumentację i materiały, prowadzimy szkolenia oraz publikujemy wyniki prowadzonych prac badawczych.

Kontakt: Mariusz Sterzel, m.sterzel [at] cyfronet.pl



Laboratorium Algorytmów Równoległych

Laboratorium skupia swoją działalność na szeroko pojętym aspekcie obliczeniowym równoległych algorytmów ze szczególnym uwzględnieniem algorytmów uczenia maszynowego, a w szczególności możliwości ich efektywnego wykorzystania na dużych klastrach obliczeniowych.

Ze względu na rosnącą z każdym dniem ilość dostępnych danych, tradycyjne algorytmy uczenia maszynowego stają się niewystarczające, zaś szeregowy paradygmat przetwarzania jest obliczeniowo nieefektywny. Aby sprostać nowym wyzwaniom związanym z rosnącą ilością danych, w wielu przypadkach konieczne jest wykorzystanie dużych klastrów obliczeniowych, co wymusza dostosowanie stosowanych algorytmów do pracy w trybie równoległym. W Laboratorium Algorytmów Równoległych rozpatrujemy zarówno teoretyczne, jak i praktyczne aspekty związane z tym zadaniem. W szczególności skupiamy się na następujących dziedzinach: wizji komputerowej, obliczeniach tensorowych, sieciach głębokich, przetwarzaniu obrazów o niskiej jakości, rozpoznawaniu obrazów podwodnych, klasyfikacji danych hiperspektralnych, klasyfikacji danych histopatologicznych, niezbalansowaniu danych.

W obrębie rozpatrywanych dziedzin rozwijamy nowe algorytmy wykorzystujące obliczenia w trybie równoległym, w szczególności badamy aspekty teoretyczne oraz praktyczne związane z tym zjawiskiem.

Kontakt: Bogusław Cyganek, cyganek [at] agh.edu.pl

Laboratorium Metod Informatycznych w Medycynie

Główne zadania Laboratorium skupiają się w dwóch sferach. Pierwszą z nich jest działalność badawcza, obejmująca gruntowną analizę i weryfikację dostępnych oraz potencjalnych odpowiedzi na wyzwania znajdujące się na styku medycyny i technologii informatycznych. Druga obejmuje projektowanie, opracowywanie i późniejszą eksploatację dedykowanych aplikacji i platform dla zastosowań w medycynie. Ten zakres obejmuje również monitorowanie stanu bezpieczeństwa opracowywanego oprogramowania oraz mechanizmów składowania i przetwarzania danych.

Dzięki kompleksowemu ujęciu procesów: od rozpoznania problemu badawczego, poprzez analizę potrzeb użytkowników, aż po ostateczne wdrożenie, Laboratorium skutecznie realizuje swoją misję wsparcia środowiska naukowego i medycznego. W ramach upowszechniania wiedzy eksperckiej, członkowie Zespołu publikują wyniki badań w czasopismach naukowych, biorą udział w przygotowywaniu materiałów informacyjnych i prowadzą konsultacje dla użytkowników.

Pracownicy Laboratorium nawiązują współpracę z renomowanymi krajowymi i zagranicznymi instytutami badawczymi i ośrodkami informatyki medycznej. Efektami są m.in. trwające oraz już zrealizowane projekty z istotnym udziałem członków Zespołu:

- Sano: Centrum Doskonałości Nowych Metod Diagnostyki Obliczeniowej i Spersonalizowanej Terapii,
- PRIMAGE: Wykorzystanie metod symulacji komputerowej i analizy obrazów w spersonalizowanej diagnostyce raka,
- Virolab: A Virtual Laboratory for Decision Support in Viral Disease Treatment,
- Gliomed: Diagnostyka glejaków na podstawie wolnokrążącego DNA guza,



- Eurvalve: Wspomaganie spersonalizowanej terapii dla przypadków choroby zastawek serca (Personalised Decision Support for Heart Valve Disease),
- CECM: Centrum Doskonałości Nowych Metod Diagnostyki Obliczeniowej i Spersonalizowanej Terapii (A Centre for New Methods in Computational Diagnostics and Personalised Therapy).

Wcześniejsza działalność obecnego zespołu Laboratorium jest przedstawiona szczegółowo na stronie: <http://dice.cyfronet.pl>.

Kontakt: Marian Bubak, bubak [at] agh.edu.pl



Laboratorium Przetwarzania Danych

Laboratorium projektuje i wdraża dedykowane aplikacje i platformy software'owe dla zastosowań w różnych dziedzinach nauki. Laboratorium tworzą specjaliści w zakresie architektur oprogramowania, programowania Front-end i Back-end, projektowania interfejsów użytkownika i user experience, DevOps, testowania, analityki wymagań. Kierunkami specjalizacji zespołu są przede wszystkim:

- opracowywanie nowatorskich metod pozyskiwania wiedzy z dostępnych danych,
- rozwój technologii wspierających przetwarzanie danych otwartych,
- integracja systemów przetwarzania danych i wiedzy z istniejącymi repozytoriami i e-infrastrukturami.

Zespół Laboratorium opracował autorskie środowisko programistyczne **InSilicoLab**, obejmujące zestaw zaawansowanych narzędzi i bibliotek programistycznych pozwalających na budowę i rozwój dedykowanych portali naukowych. Portale bazujące na InSilicoLab są zaprojektowane w taki sposób, aby gromadzić w jednym miejscu całość narzędzi, których badacze potrzebują dla obliczeń in silico. Zaletami są przede wszystkim:

- łatwe uruchamianie nawet skomplikowanych, długich i wymagających wielu obliczeń, eksperymentów użytkownika,
- możliwość wygodnego opisu, kategoryzacji i wyszukiwania danych wejściowych lub wyników.

Technologię InSilicoLab wyróżnia dążenie do jak największej użyteczności budowanych za pomocą środowiska narzędzi. Ta sfera obejmuje zarówno użyteczność dla rozwiązywania problemów

naukowych w danej dziedzinie, jak również przyjazność portalu dla jego końcowego odbiorcy. Laboratorium nawiązuje współpracę z renomowanymi jednostkami naukowymi i badawczymi w ramach polskich i międzynarodowych projektów. Efektami współpracy są między innymi:

- **Rozwój Platformy IS-EPOS** (<https://tcs.ah-epos.eu/>) w ramach serii projektów związanych z European Plate Observing System - EPOS (<https://www.epos-eu.org/>). Portal oraz zorganizowane wokół niego narzędzia są zorientowane na badanie i analizę sejsmiczności oraz innych zjawisk spowodowanych działalnością człowieka (np. eksploatacja zasobów w ramach kopalni, tworzenie sztucznych zbiorników wodnych). Portal jest zintegrowany z europejską infrastrukturą EPOS.
- **Budowa Portalu Sat4Envi** (<https://dane.sat4envi.imgw.pl/>), udostępniającego dane satelitarne z programu Copernicus. Portal umożliwia wyszukiwanie, oglądanie, zamawianie i pobieranie danych satelitarnych oraz ich produktów pochodnych z poziomu przeglądarki internetowej.
- **Opracowanie i utrzymanie Portalu EOSC** (<https://eosc-portal.eu/>) w ramach serii projektów związanych z European Open Science Cloud (EOSC). Portal zapewnia dostęp do zasobów wielu europejskich e-infrastruktur i infrastruktur badawczych za pomocą zunifikowanego systemu uwierzytelniania użytkownika. Działania EOSC skupiają się wokół realizacji paradygmatu Open Science (otwartej nauki).

Zespół Laboratorium był również zaangażowany w tworzenie **Portalu PLGrid** (<https://portal.plgrid.pl/>), udostępniającego naukowcom wiele pakietów oprogramowania, bibliotek i narzędzi naukowych.

Kontakt: Tomasz Szepieniec, t.szepieniec [at] cyfronet.pl

Laboratorium Technologii Chmurowych

Laboratorium zajmuje się projektowaniem i eksploatacją chmury dla nauki, a także narzędzi służących efektywnemu korzystaniu z niej. Zespół opracowuje kompleksowe środowiska dostępu do rozproszonych danych, uwzględniające zarówno kwestie bezpiecznego składowania i przetwarzania danych w chmurze, jak i wygodne interfejsy dostępowe (portale, aplikacje) dla użytkownika końcowego.



Mając na uwadze dynamiczny rozwój nowych technologii przetwarzania i składowania danych w chmurze, Laboratorium w trybie ciągłym prowadzi prace badawcze i publikuje wyniki. Korzystając z wiedzy eksperckiej zespołu, aktywnie wspiera inicjatywy naukowe, w tym międzynarodowe projekty i e-infrastruktury.

Flagowym produktem Laboratorium jest **Onedata**: globalnie skalowalny system zarządzania danymi, ujednolicający dostęp do danych przechowywanych w rozproszonych systemach. Onedata równie dobrze dpowiada na potrzeby zarówno

małych grup użytkowników, jak i dużych, międzynarodowych środowisk naukowych. System umożliwia użytkownikom korzystanie z jednorodnego systemu zarządzania danymi zarówno do przechowywania danych osobistych jak i związanych z pracą zawodową, np. wyników badań, oraz daje możliwość dostępu do nich w wydajny sposób z dowolnego urządzenia.

Onedata jest aktualnie uruchomione w ramach kilku infrastruktur europejskich, w tym w projektach: PLGrid, INDIGO-DataCloud, EGI Data Hub oraz Human Brain Project. W projekcie HBP Onedata wykazało możliwość dostarczenia systemu zarządzania danymi pozwalającego na spełnienie wymagań przepustowości dla aplikacji wizualizacji przekrojów mózgu w czasie rzeczywistym.

Więcej informacji na stronie: <https://onedata.org>.

Kontakt: Łukasz Dutka, l.dutka [at] cyfronet.pl

Laboratorium Zastosowań Technik Obliczeniowych

Kluczowym aspektem działania Laboratorium jest opracowywanie modeli numerycznych oraz symulacji komputerowych w oparciu o architektury superkomputerowe, ze szczególnym uwzględnieniem wykorzystania praktycznego w przemyśle. Laboratorium prowadzi własne badania dotyczące technik obliczeniowych, jak również śledzi i analizuje nowe rozwiązania w ich zastosowaniu.

Zespół Laboratorium nawiązuje współpracę w ramach polskich i międzynarodowych konsorcjów, wspólnie z partnerami realizując projekty ukierunkowane na opracowanie lub udoskonalenie procesów produkcji wyrobów metalowych. Wśród najważniejszych efektów tych działań znajdują się:

- **VirtROLL – Virtual strip rolling mill:** głównym celem projektu było stworzenie systemu komputerowego wspomagającego elastyczne projektowanie technologii walcowania wyrobów płaskich na podstawie wyników obliczeń dedykowanych modułów do symulacji numerycznych.
- **PROTEUS-RS - Strategie optymalizacji procesów wytwarzania produktów długich w celu podniesienia jakości wyrobu gotowego poprzez minimalizację naprężeń rezydualnych:** projekt przewiduje wykonanie modeli oraz szeregu symulacji numerycznych w oparciu



o architektury komputerowe wysokiej wydajności na potrzeby projektowania procesu wytwarzania metalowych elementów długich.

Kontakt: Łukasz Rauch, lrauch [at] agh.edu.pl

Laboratorium Technik Wizualnych

Laboratorium Technik Wizualnych (Media Lab 5.0) jest studiem filmowo-telewizyjnym, którego głównym zadaniem jest produkcja materiałów dokumentujących i promujących osiągnięcia ACK Cyfronet AGH.

Realizowane filmy i programy telewizyjne popularno-naukowe i szkoleniowe pokazują efekty prac badawczych użytkowników infrastruktury obliczeniowej Cyfronetu.

W Laboratorium prowadzone są warsztaty dziennikarskie, sztuki operatorskiej, realizacji dźwięku, sztuki montażu, grafiki komputerowej i obsługi studia telewizyjnego. To również miejsce prowadzenia prac badawczych z technologii wizyjnych, akustyki oraz technik transmisji on-line.

Laboratorium to studio filmowo-telewizyjne o powierzchni około 100 m² wyposażone w ruszt z oświetleniem, nagłośnioną, ze sztucznym horyzontem i scenografią. Studio połączone jest z salą audytoryjną na około 50 miejsc. Studio skomunikowane jest z reżyserkami: WIZJI i ŚWIATŁA (ze stanowiskami: regulacji torów kamerowych, realizacji wizji, nagrywania i odtwarzania, grafiki komputerowej, światła), DŹWIĘKU z konsolą Yamaha LS 9 oraz LEKTORKĄ. W Laboratorium znajdują się dwie MONTAŻOWNIE z zestawami komputerowymi z oprogramowaniem Adobe CS6. Laboratorium posiada kamery Panasonic AG-HPX 371 (3 szt.) oraz inny sprzęt filmowo-telewizyjny do realizacji obrazu, światła i dźwięku w produkcjach studyjnych i plenerowych.

Studio współpracuje z profesjonalnymi twórcami i realizatorami oraz studentami krakowskich uczelni.

Filmy zrealizowane przez MediaLab 5.0 można obejrzeć między innymi:

- na kanale ACK Cyfronet AGH w serwisie YouTube (<https://www.youtube.com/user/CyfronetAGH/videos>),
- na platformie Pionier.tv (<https://pionier.tv/wideo-tag/cyfronet/>).

Kontakt: Jacek Przybylski, j.przybylski [at] cyfronet.pl



Laboratorium Akceleracji Obliczeń i Sztucznej Inteligencji

Nasza misja

Dzięki kompetencjom w dziedzinie algorytmów sztucznej inteligencji (SI) i znajomości nowoczesnych technik obliczeniowych wspieramy naukowców w ich pracach badawczych. Nasza wiedza i doświadczenie pozwala na skuteczną realizację uczenia maszynowego przy wykorzystaniu dostępnej w ACK Cyfronet AGH, dedykowanej do obliczeń neuronowych, partycji klastra komputerowego Prometheus. Partycja do obliczeń SI zbudowana została w oparciu o cztery wydajne serwery wyposażone w osiem kart GPGPU nVidia Volta V100 każdy. Całkowita moc obliczeniowa partycji to ponad 4 PetaFlops, czyli ponad cztery biliardy (4×10^{15}) obliczeń dla SI na sekundę.

Specjalizacje domenowe

Poniżej przedstawiono wybrane zastosowania sztucznej inteligencji w przetwarzaniu języka naturalnego, przetwarzaniu obrazów i analizie szeregów czasowych w różnych problemach badawczych i zastosowaniach praktycznych.

Przetwarzanie języka naturalnego

Przetwarzanie języka naturalnego to zagadnienie, w którym Laboratorium posiada wieloletnie doświadczenie. Od wielu lat tworzymy i rozwijamy narzędzie umożliwiające przeszukiwanie, porównywanie oraz klasyfikację dokumentów tekstowych. Praktycznym owocem tych prac jest usługa Scholar, dostępna w ACK Cyfronet AGH na platformie usług PLGrid. Jednym z ważnych zagadnień podejmowanych badawczo jest analiza efektywności metod redukcji dokładności reprezentacji danych tekstowych i ich wpływu na wyniki działania algorytmów. Udało nam się opracować modyfikacje pozwalające na 10-krotne zmniejszenie zużycia energii obliczeń w stosunku do oryginalnej implementacji.

Pojawienie się rozwiązań opartych o sieci neuronowe zrewolucjonizowało dziedzinę. Zajmujemy się badaniami nad możliwością kompresji i implementacji sprzętowej sieci do oceny sentymentu, które pokazały, że możliwa jest redukcja dokładności reprezentacji wag sieci do 8 a nawet 4 bitów przy zachowaniu niemal niezmienionej skuteczności działania sieci. Dodatkowo zajmujemy się zagadnieniami semi-nadzorowanego uczenia, w których ilość dostępnych, oznaczonych danych była bardzo ograniczona, a docelowe kategorie zmieniały się w trakcie działania systemu. Nasze badania pokazały, że możliwe jest opracowanie rozwiązania osiągającego dokładność nawet 98.9%.

Przetwarzanie obrazów

Przetwarzanie obrazów przy pomocy sieci neuronowych skupia się wokół prac Laboratorium, które związane są z rozpoznawaniem i detekcją obiektów na potrzeby aplikacji medycznych. W ich trakcie opracowany został system do klasyfikacji zmian nowotworowych w pobranych od

zwierząt próbkach z badania cytologicznego. Zaproponowany system osiągnął skuteczność około 96% dla wybranych trzech typów nowotworów. W pracach zostały wykorzystane modele głębokich sieci typu Resnet-50 oraz Resnet-152 i opracowany został specjalny schemat trenowania z wykorzystaniem algorytmów genetycznych dedykowany do klasyfikacji, który pozwala na selektywne wybieranie obszaru percepcji dla operacji klasyfikacji. Jednym z najnowszych osiągnięć w tym obszarze jest model do detekcji bazujący na architekturze Yolo3, który osiąga bardzo dobrą wartość współczynnika mAP, równą 0.86 dla specjalnie przygotowanych obrazów wykonanych z wykorzystaniem wielu preparatów cytologicznych niskiej jakości. Opracowywany system ma być docelowo wykorzystywany w codziennej pracy lekarzy weterynarii.

VetDemo

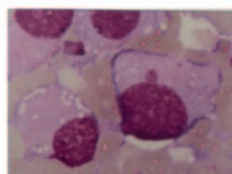
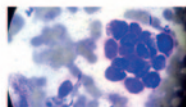
[New order](#) [My orders](#) [Log out](#)

Order

ID: 729d8955-9b38-4a8b-9135-942e803fa15e

Created: July 4, 2018, 2:08 p.m.

Finished: (Status)



VetDemo

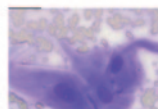
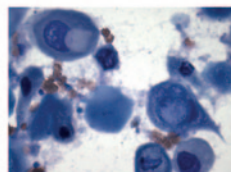
[New order](#) [My orders](#) [Log out](#)

Order

ID: 520fbb64-5b6c-40de-9764-3f9bd75e5c70

Created: July 4, 2018, 2:10 p.m.

Finished: (Status)



Status:

SCC

Status:

SCC

Status:

SCC

Status:

SCC

Status:

SCC

Status:

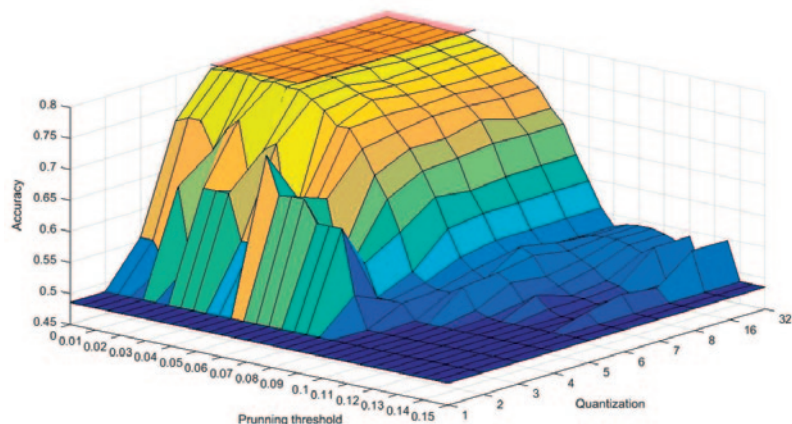
SCC

Zdjęcia preparatów cytologicznych przedstawiające zmiany nowotworowe skóry psa

Analiza szeregów czasowych

Prace Laboratorium dotyczą także modelowania szeregów czasowych służącego detekcji sytuacji nietypowych. Szczegółowe prace obejmują aplikacje praktyczne, takie jak predykcja anomalii w celu ochrony przed katastrofalnym uszkodzeniem magnesów Wielkiego Zderzacza Hadronów (LHC) oraz innych urządzeń stowarzyszonych znajdujących się w CERN. Opracowany system do detekcji quenchy w magnesach nadprzewodzących wykorzystuje sieci neuronowe GRU oraz LSTM oraz dedykowany post-processing, który pozwolił dodatkowo na klasyfikację wykrytych anomalii. Co więcej, opracowany algorytm został zrealizowany sprzętowo w architekturze FPGA Xilinx Zynq UltraScale+ MPSoC 285, co pozwoliło na uzyskanie bardzo niskiej latencji odpowiedzi.

W obszarze medycyny podejmowany jest problem detekcji zaskórnic u osób hospitalizowanych, które długi czas są unieruchomione w łóżku szpitalnym. Analizy dotyczą znanego problemu zaniku funkcji wspomagania układu krążenia przez pracę mięśni nóg i są realizowane we współpracy z Medical University of Graz.



Konferencja Użytkowników Komputerów Dużej Mocy (KU KDM)



Podstawowym założeniem konferencji KU KDM było zainicjowanie corocznych spotkań naukowych gromadzących użytkowników, którzy wykonują obliczenia z wykorzystaniem komputerów dużej mocy i oprogramowania, udostępnianych przez ACK Cyfronet AGH.

Po raz pierwszy konferencja została zorganizowana w 2008 roku, a jej program obejmował liczne wystąpienia naukowców – prezentujące wyniki prac badawczych uzyskanych przy wykorzystaniu zasobów sprzętowych i oprogramowania oferowanego w Cyfronecie. Konferencję uświetniły wykłady zaproszonych gości – Norberta Attiga z Jülich Supercomputing Centre i Jaapa A. Kaandorpa z University of Amsterdam. Pierwsza edycja konferencji spotkała się z dużym uznaniem wśród jej uczestników. Tego typu wydarzenie było bardzo oczekiwane i potrzebne.

Obecna tematyka konferencji KU KDM koncentruje się na obliczeniach i symulacjach dużej skali, nowych algorytmach w informatyce, narzędziach i technikach w systemach dużej mocy obliczeniowej, zastosowaniach informatyki w dydaktyce oraz bazach danych. Jednakże jej głównym celem jest możliwość prezentacji osiągnięć naukowych szerokiemu gronu badaczy

z różnych dziedzin. Jest to również okazja do przybliżenia użytkownikom samego Centrum oraz jego zasobów, w tym infrastruktury PLGrid.

Użytkownicy, którzy na co dzień prowadzą badania z wykorzystaniem zasobów Cyfronetu, uzyskują możliwość przedstawienia roli tych zasobów, typowych scenariuszy użycia i aspektów wydajnościowych w trakcie swoich prezentacji. Uczestniczący w konferencji przedstawiciele Cyfronetu, dzięki sposobności spotkania się z tymi naukowcami, prowadzenia rozmów i konsultacji, pozyskują wiedzę niezbędną do podjęcia działań zmierzających do dostosowania infrastruktury obliczeniowej do potrzeb wymagających środowisk naukowych. Atmosfera sprzyja również wy-

mianie doświadczeń pomiędzy naukowcami reprezentującymi te same lub różne dyscypliny oraz zapoznaniu się z nowymi technologiami oraz domenowo-specyficznymi usługami wdrażanymi przez Centrum.

Do ważnych części konferencji należą spotkania z dostawcami sprzętu i oprogramowania dla Cyfronetu, a także dyskusja panelowa na temat efektywnego wykorzystania tych zasobów. Ta ostatnia cieszy się zawsze bardzo dużym zainteresowaniem użytkowników – naukowców, którzy mogą wtedy uzyskać informacje o nowościach w infrastrukturze obliczeniowej Centrum, a także bezpośrednio poinformować ekspertów z Cyfronetu o problemach, jakie napotkali w trakcie pracy z tą infrastrukturą.

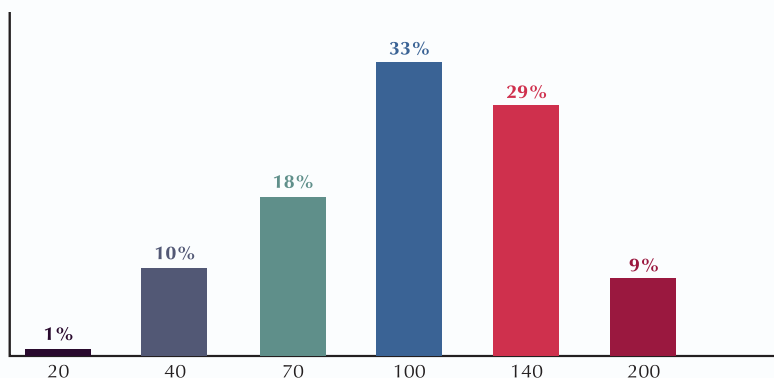


W roku 2020, z powodu trwającej pandemii koronawirusa (COVID-19), konferencja została zorganizowana online. Więcej informacji: <https://www.cyfronet.pl/doack2020/>

Artykuły naukowe opracowane na podstawie najlepszych prezentacji konferencyjnych są publikowane w jednym z dwóch znanych czasopism naukowych z dziedziny IT: Computing and Informatics (CAI) (<http://www.cai.sk>) lub Computer Science (CSCI) (www.csci.agh.edu.pl).



Punktacja artykułów naukowych według MEiN, powstałych w 2020 roku dzięki zasobom ACK Cyfronet AGH





Sat4Envi – kosmiczne technologie, powszechne wykorzystanie

Dzięki udostępnieniu danych satelitarnych z programu Copernicus, projekt Sat4Envi otworzył nowy rozdział dla podejmowania optymalnych decyzji dotyczących planowania przestrzennego i szybkiego reagowania w przypadku wystąpienia zagrożeń środowiskowych.

Dane satelitarne na wyciągnięcie ręki

Pod adresem *dane.sat4envi.imgw.pl* dostępne jest kluczowe dokonanie projektu – System Obsługi Klienta, czyli obsługiwane z poziomu przeglądarki internetowej narzędzie dostępu do hurtowni danych satelitarnych Sat4Envi. Dzięki tej platformie użytkownik końcowy zyskuje możliwość wyszukiwania, oglądania, zamawiania i pobierania danych satelitarnych oraz ich produktów pochodnych. System służy przede wszystkim do monitorowania powierzchni ziemi, monitorowania atmosfery i pogody oraz rozpoznawania zagrożeń meteorologicznych. Dostęp do wszystkich danych jest całkowicie darmowy i nie zależy od afiliacji, czy wykonywanej profesji. W wersji podstawowej nie wymaga rejestracji. Po zalogowaniu użytkownik ma możliwość skorzystania z dodatkowych funkcjonalności, jak chociażby zapisywanie „koszyka” najczęściej wybieranych pomiarów, czy dostęp do metadanych, jakimi opisane są prezentowane obrazy satelitarne. Dzięki nałożeniu warstw WMS (Web Map Service), zawierających m.in. granice jednostek administracyjnych, narzędzie jest funkcjonalne dla jednostek administracji publicznej czy instytucji regionalnych. Samo działanie Systemu Obsługi Klienta jest intuicyjne i pozwala na szybkie dotarcie do pożądanego typu danych.

Projekt i partnerzy

System operacyjnego gromadzenia, udostępniania i promocji cyfrowej informacji satelitarnej o środowisku (Sat4Envi) ma na celu szerokie i wygodne udostępnienie satelitarnych danych programu Copernicus oraz danych pochodzących z innych satelitów środowiskowych i meteorologicznych.

Za sukcesem Sat4Envi stoi zaangażowanie wielu specjalistów z czterech ściśle współpracujących jednostek. Liderem jest Instytut Meteorologii i Gospodarki Wodnej – Państwowy Instytut Badawczy, a partnerami w projekcie: Centrum Badań Kosmicznych Polskiej Akademii Nauk, Polska Agencja Kosmiczna oraz Akademickie Centrum Komputerowe CYFRONET AGH. IMGW-PIB odpowiada przede wszystkim za rozbudowę stacji odbioru i przetwarzania danych z satelitów okołobiegunowych (stacja powstała w Krakowie) oraz pełni rolę koordynatora. Wspólny wysiłek pozwala sprawnie realizować pozostałe założenia projektu, którymi są:

- budowa nowoczesnego archiwum danych satelitarnych (IMGW-PIB, CYFRONET),
- centrum szkoleniowe nowych technologii satelitarnych (CBK, PAK, IMGW-PIB),
- centrum udostępniania informacji naukowej dla użytkowników (CYFRONET, IMGW-PIB, CBK).

Projekt, o łącznej wartości 17 903 900 zł, jest współfinansowany ze środków Europejskiego Funduszu Rozwoju Regionalnego w ramach Programu Operacyjnego Polska Cyfrowa.

Technologiczne wyzwania

Budowa portalu z interfejsem dostępowym wymagała bardzo starannego i wieloaspektowego projektowania, uwzględniającego łatwość wyszukiwania danych czy szybkość odnajdywania określonych funkcji. Przyjmując za cel stworzenie produktu o wysokiej funkcjonalności również dla użytkowników „niezaawansowanych”, zespół programistyczny rozpoczął działania od rozbudowanej oceny wymagań. Analizowano przypadki użycia w różnych grupach użytkowników docelowych i przeprowadzano wielokrotne weryfikacje z użyciem makiet graficznych i prototypowych implementacji. W pracach nad interfejsem specjaliści w zakresie user experience sięgnęli również do techniki wywiadów pogłębionych w grupach reprezentantów przyszłych użytkowników (jak IMGW, CBK, PAK, Pracownia Teledetekcji Morza PAN czy Pieniński Park Narodowy).

- Rolą Cyfronetu w projekcie była realizacja zadań programistycznych związanych z budową i utrzymaniem Systemu Obsługi Klienta, ale też zapewnienie infrastruktury informatycznej niezbędnej do gromadzenia i udostępniania danych. Cyfronet jest operatorem połączenia internetowego z IMGW-PIB Kraków, a w ramach infrastruktury PLGrid udostępnia zasoby obliczeniowe niezbędne dla prawidłowego funkcjonowania Sat4Envi. Z kolei do przetwarzania danych i tworzenia informacji z surowych danych satelitarnych wykorzystywany jest zainstalowany w Cyfronecie superkomputer Prometheus – podkreśla prof. Kazimierz Wiatr, Dyrektor ACK Cyfronet AGH.

Sat4Envi wspiera w sytuacjach kryzysowych

Analiza danych satelitarnych z jak najmniejszymi opóźnieniami w stosunku do ich pozyskania może mieć kluczowe znaczenie dla prognozowania, monitorowania oraz przeciwdziałania skutkom kataklizmów. W szczególności mówimy tu o monitorowaniu powodzi, ocenie skali pożarów, czy monitorowaniu szkód w uprawach rolnych powstałych w wyniku suszy lub przemarzania. Ze względu na ważność tych działań i konieczność szybkiego reagowania, projekt Sat4Envi uwzględnia wsparcie dla specjalnych grup użytkowników, takich jak Centrum Informacji Kryzysowej.

Więcej informacji na temat projektu znajduje się na stronie: <https://sat4envi.imgw.pl/>.



Prometheus i jego infrastruktura służą udostępnianiu danych z programu Copernicus. Źródło: Sat4Envi



EPOS - multidyscyplinarna platforma wsparcia badań w zakresie nauk o Ziemi

Efektywne monitorowanie procesów zachodzących we wnętrzu Ziemi oraz ich skutków odczuwalnych na powierzchni wymaga wykorzystania różnorodnych narzędzi i metodologii. EPOS – European Plate Observing System (System Obserwacji Płyty Europejskiej) ma w tym zakresie dostarczyć badaczom kompleksowych rozwiązań, przekraczających zarówno granice państw, jak i pojedynczych dyscyplin naukowych.

W październiku 2018 roku EPOS otrzymał status ERIC: European Research Infrastructure Consortium (Europejskiego Konsorcjum Infrastruktury Badawczej). W ten sposób stworzone zostały ramy prawne i organizacyjne pozwalające na występowanie na arenie międzynarodowej już nie jako projekt współrealizowany przez wielu partnerów, a jako odrębny podmiot z główną siedzibą w Rzymie.

Ideą stojącą za EPOS jest lepsze poznanie funkcjonowania Ziemi jako złożonego systemu, w którym z jednej strony zdarzenia naturalne (jak erupcje wulkanów, powodzie czy trzęsienia ziemi) wpływają na społeczeństwo i gospodarkę, a z drugiej, działalność człowieka zmienia środowisko. Z tego względu do efektywnych badań potrzebne jest podejście wieloaspektowe. Poprzez integrację europejskich infrastruktur badawczych w ponadnarodowy system udostępniający dane, ich produkty, oprogramowanie oraz usługi, EPOS opracowuje dostęp do możliwości, które do tej pory występowały wyłącznie w formie rozproszonej. Inicjatywa skupia się na dziesięciu głównych, powiązanych ze sobą węzłach tematycznych (Thematic Core Services – TCS):



- Dane sejsmologiczne
- Obserwatoria przy uskokach tektonicznych
- Dane i produkty GNSS (Global Navigation Satellite Systems)
- Obserwacje wulkanów
- Dane satelitarne
- Badania geomagnetyzmu
- Zagrożenia antropogeniczne
- Informacje i modelowanie geologiczne
- Laboratoria Wielkoskalowe
- Stanowiska badań geoenergetycznych związane z energią nisko-węglową.

Źródło: EPOS - European Plate Observing System

Obliczeniowe nauki o Ziemi

Udostępnienie infrastruktury badawczej jako platformy internetowej wymaga ścisłej współpracy specjalistów w zakresie technologii informatycznych z przedstawicielami

nauk o Ziemi. Synergiczne działania są realizowane na wielu poziomach – począwszy od prowadzonych przez naukowców pomiarów i eksperymentów generujących wielką ilość surowych danych, które są składowane i przetwarzane w systemach informatycznych – między innymi w ramach infrastruktury Cyfronetu. Współpraca jest też bardzo istotna przy budowie narzędzi służących do wizualizacji danych, sterowanych z poziomu przeglądarki internetowej za pomocą interfejsów graficznych (GUI). Docelowo, EPOS ma połączyć środowiska badawcze w ramach jednej, multidyscyplinarnej platformy, dostarczyć efektywnych i bezpiecznych narzędzi dostępowych i analitycznych oraz zainicjować powstawanie nowych usług w oparciu o dotychczasowe osiągnięcia. Co ważne, dane geofizyczne i geologiczne, podobnie jak zaawansowane oprogramowanie do ich wizualizacji i analizy, będą dostępne w sposób otwarty i bezpłatny. Dzięki temu EPOS wspiera interdyscyplinarne badania dotyczące przyczyn i skutków procesów zachodzących wewnątrz i na powierzchni Ziemi. Ma to istotne znaczenie w świetle monitorowania zagrożeń środowiskowych, takich jak: trzęsienia ziemi, powódzie, osuwiska czy wybuchy wulkanów, jak również zagrożeń antropogenicznych związanych m.in. z działalnością kopalni. Poprzez kompleksową analizę zjawisk już zaobserwowanych, można lepiej przygotować się na zdarzenia, które jeszcze nie miały miejsca, starając się jak najbardziej ograniczyć ich negatywne skutki.

Platforma EPOS w fazie testów operacyjnych jest dostępna pod adresem: <https://www.ics-c.epos-eu.org/>, a portal integrujący informacje o działaniach EPOS pod adresem: <https://www.epos-eu.org/>.

Działania Cyfronetu na rzecz EPOS

ACK Cyfronet AGH uczestniczy w kolejnych projektach budujących EPOS, wspierając inicjatywę zarówno zasobami sprzętowymi, jak i wiedzą oraz umiejętnościami specjalistów. Równolegle prowadzone działania obejmują kilka sfer, jedną z których jest opracowanie platformy cyfrowej dla węzła tematycznego Zagrożeń Antropogenicznych (TCS AH).

Platforma udostępniona pod adresem: <https://tcs.ah-epos.eu/> stanowi narzędzie do analizy sejsmiczności antropogenicznej i związanych z nią zagrożeń oraz do oceny potencjalnego wpływu eksploatacji geo-zasobów na środowisko. Dzięki danym pochodzącym ze stacji sejsmicznych oraz udostępnionym informacjom przemysłowym, platforma pozwala łatwiej analizować takie procesy jak: zalewanie sztucznych zbiorników, wydobywanie surowców, gazów łupkowych czy wód podziemnych. Oprócz działań związanych z udostępnianiem i wizualizacją danych, specjaliści z Cyfronetu dbają o mechanizmy uwierzytelniania użytkowników (m.in. poprzez dostosowanie systemu Unity) oraz skalowanie zasobów dla zwiększającej się liczby użytkowników. Dodatkowo, Cyfronet zapewnia wirtualną przestrzeń roboczą do uruchamiania aplikacji od różnych dostawców oraz porządkowania danych. Obecnie platforma TCS AH liczy ponad 1200 zarejestrowanych użytkowników.

Z kolei w ramach projektu EPOS SP (EPOS Sustainability Phase) Cyfronet jest zaangażowany w działania mające na celu pozyskanie nowych użytkowników infrastruktury EPOS, w tym partnerów przemysłowych.

Równolegle trwają prace związane z utrzymywaniem zgodności z serwisami centralnymi EPOS oraz integracją kolejnych pakietów oprogramowania wykorzystywanego w badaniach naukowych.

Dodatkowo, uczestnicząc w projektach krajowych, Cyfronet działa na rzecz rozbudowy polskiej Infrastruktury Badawczej EPOS-PL. Najnowsze prace, skupione w projekcie EPOS PL+, dotyczą opracowania oprogramowania wspierającego planowanie i prowadzenie badań z użyciem metod sztucznej inteligencji na danych IB EPOS-PL. Powstała platforma będzie umożliwiała łatwe korzystanie z danych zintegrowanych w programie, dobór metod uczenia maszynowego oraz uruchamianie zasobochłonnych etapów przetwarzania na zasobach obliczeniowych polskich Centrów Komputerów Dużej Mocy.

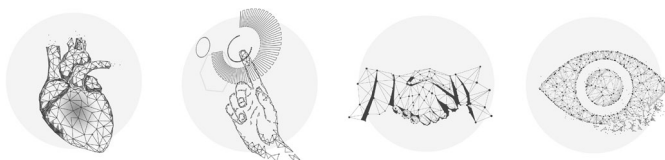


Sano: Centrum Zindywidualizowanej Medycyny Obliczeniowej – Międzynarodowa Fundacja Badawcza

W latach 2019-2026 Akademickie Centrum Komputerowe Cyfronet AGH wraz z pięcioma partnerami realizuje wyjątkowy projekt w ramach programu *Teaming for Excellence* Horyzontu 2020 Unii Europejskiej, programu Międzynarodowych Agend Badawczych prowadzonego przez Fundację na rzecz Nauki Polskiej oraz przy wsparciu finansowym Ministerstwa Nauki i Szkolnictwa Wyższego. W wyniku tego projektu w Krakowie powstało **Sano** – Centrum Zindywidualizowanej Medycyny Obliczeniowej.

Misja **Sano** obejmuje:

- rozwój nowych metod obliczeniowych, algorytmów, modeli i technologii związanych z medycyną spersonalizowaną,
- wprowadzenie do codziennej praktyki opieki zdrowotnej nowych rozwiązań diagnostycznych i terapeutycznych opartych na symulacjach komputerowych,
- stymulowanie powstawania i rozwoju przedsiębiorstw tworzących technologie umożliwiające wprowadzanie nowych metod diagnostycznych i terapeutycznych,
- wkład w nowe programy kształcenia dla potrzeb nowoczesnej, zindywidualizowanej medycyny.



Centrum Sano (<https://sano.science/>) zlokalizowane jest w Krakowie, gdzie kształcą się liczni wysokiej klasy specjaliści w zakresie nauk medycznych i informatyki. Co ważne, krakowskie szpitale badawcze są wysoko oceniane w środowisku naukowym i systematycznie rośnie w tym mieście liczba firm w sektorze *life science*.

Rozwój **Centrum Sano** przyczyni się bezpośrednio do ulepszenia poziomu nauki w naszym regionie, poprzez nowe kontakty naukowe oraz możliwość kształcenia doktorantów na światowym poziomie, a ponadto wzmocni transfer wiedzy i technologii do gospodarki sprzyjając powstawaniu nowych firm działających w obszarze zaawansowanych technologii. Oddziaływanie Centrum będzie miało charakter ponadregionalny, czy wręcz międzynarodowy, przyczyniając się do rozwoju nowoczesnej medycyny, a w konsekwencji do wzrostu poziomu opieki zdrowotnej.

Program działania Centrum jest oparty m.in. na Polskiej Strategii Inteligentnej Specjalizacji i ma na celu rozwój współpracy nauki i biznesu w skali międzynarodowej. Kluczowymi wskaźnikami sukcesu są: liczba wysoko cytowanych publikacji naukowych i grantów uzyskanych przez Centrum, liczba rozwiązań opartych na modelach obliczeniowych wprowadzonych do praktyk klinicznych oraz liczba innowacyjnych rynkowych produktów i usług.

Centrum Zindywidualizowanej Medycyny Obliczeniowej Sano powstało dzięki międzynarodowej współpracy następujących instytucji: ACK Cyfronet AGH, Klaster LifeScience Kraków – Krajowy Klaster Kluczowy, University of Sheffield i Insigneo Institute, Forschungszentrum Jülich, Fraunhofer Institute for Systems and Innovation Research ISI oraz Narodowe Centrum Badań i Rozwoju.

EOSC – ponadnarodowa integracja zasobów naukowych



European Open Science Cloud (EOSC) jest inicjatywą zmierzającą do stworzenia wirtualnego środowiska wypełniającego założenia paradygmatu Otwartej Nauki. EOSC ma w łatwy i przejrzysty sposób udostępnić nie tylko dane naukowe, ale też zaawansowane narzędzia i zasoby do ich składowania, udostępniania, przetwarzania i zarządzania nimi.

W ramach EOSC tworzone są połączenia pomiędzy istniejącymi e-infrastrukturami badawczymi i naukowymi. Integracja następuje między innymi poprzez ujednolicenie zasad dostępu i autoryzacji dla badaczy z różnych państw. Dzięki osiągnięciom projektu EOSC-Hub stworzono platformę realizującą powyższe zadanie. EOSC-Portal stanowi miejsce styku pomiędzy dostawcami usług i zasobów naukowych oraz badaczami, którzy mogą skorzystać nie tylko z samych zasobów, ale też wsparcia technicznego i szkoleń. Zespół ACK Cyfronet AGH odegrał w tym zakresie kluczową rolę, zostając dostawcą portalu oraz rozwijając serwis Marketplace, z rozbudowanym katalogiem usług udostępnianych przez partnerów EOSC.

Dalszy rozwój EOSC

Kluczowymi założeniami dalszego rozwoju EOSC jest rozbudowa Portalu oraz integracja usług od większej ilości dostawców.

Realizowany w latach 2019-2021 projekt EOSC Enhance (<https://www.eosc-portal.eu/enhance>) miał na celu usprawnienie Portalu pod kątem wygody i szybkości użytkowania. Prowadzone przez naszych specjalistów prace dotyczyły między innymi zaawansowanej analityki zachowań użytkowników w celu stworzenia i wdrożenia najlepszych praktyk *user experience*. Jednocześnie udostępniono nowe funkcjonalności.



Od września 2019 roku trwa projekt EOSC Synergy (<https://www.eosc-synergy.eu/>), mający na celu wdrożenie standardów EOSC dla kolejnych 9-ciu krajowych e-infrastruktur. Cyfronet, oprócz koordynacji działań na poziomie narodowym, wspiera proces planistycznie, szukając nowych, skutecznych rozwiązań dla integracji w innych krajach.



Z kolei EOSC Future jest kontynuacją i rozszerzeniem dotychczas prowadzonych prac. Projekt ma podwyższyć jakość ekosystemu EOSC dla jeszcze lepszego wsparcia europejskiej nauki oraz zachęcić naukowców do korzystania z udostępnianych zasobów. Ten cel jest realizowany poprzez dalsze prace nad Portalem, integrację kolejnych infrastruktur i inicjatyw naukowych. Dodatkowo, EOSC Future ma zapewnić długotrwałą stabilność EOSC. W ramach projektu Cyfronet odpowiada za rozwój jednego z trzech głównych komponentów Portalu: części skierowanej do użytkownika, w szczególności za jego elementy wspierające funkcjonalności Marketplace.



Równocześnie z rozwojem Portalu i towarzyszących platform w EOSC prowadzone są prace nad stworzeniem przejrzystych zasad współpracy, zarówno z dostawcami usług, jak i użytkownikami końcowymi. Tworzenie struktur formatywnych jest koordynowane przez EOSC Sekretariat, w ramach którego tworzone są dyskusyjne grupy robocze. Ich celem jest zbieranie wymagań i opracowywanie form kontaktu oraz procedur.

Grupa projektów EOSC jest finansowana z programu Unii Europejskiej „Horyzont 2020”.



EuroHPC – w stronę obliczeń w eksaskali

Europejskie Wspólne Przedsięwzięcie w dziedzinie Obliczeń Wielkiej Skali (European High-Performance Computing Joint Undertaking – EuroHPC JU) zostało powołane w celu dynamicznego rozwoju istniejącej europejskiej infrastruktury komputerów dużej mocy i tym samym udostępnienia europejskim naukowcom mocy obliczeniowych porównywalnych z tymi dostępnymi w USA, Chinach i Japonii. EuroHPC JU zrzesza 33 państwa, w tym Polskę, która dołączyła w 2018 roku, oraz prywatnych członków. Głównym celem jest zakup i rozmieszczenie dwóch superkomputerów o mocy w skali eksaflopsów (EFlops = 10^{18} operacji zmiennoprzecinkowych na sekundę), które znajdą się w światowej TOP5 (pierwszej piątce listy TOP500).

Osiągnięcie eksaskali zaplanowano na rok 2023. Wcześniej jednak uruchamiane są kolejne superkomputery o mniejszej mocy, tworzące infrastrukturę, która pozwala na przeskalowanie technologii i oprogramowania. Do czasu opracowania niniejszej publikacji, EuroHPC JU dostarczyło 2 superkomputery przed-eksaskalowe: LUMI (Finlandia) oraz Leonardo (Włochy). Trzecia maszyna tej skali (Mare Nostrum 5) zostanie uruchomiona w Hiszpanii. Dodatkowo zainstalowano lub pomyślnie uruchomiono operacyjnie 5 superkomputerów petaskalowych: Discoverer (Bułgaria), Karolina (Czechy), MeluXina (Luksemburg), Vega (Słowenia) oraz Deucalion (Portugalia).

Więcej o EuroHPC JU na stronie: <https://eurohpc-ju.europa.eu/>.



Superkomputer Karolina, Czechy. Fot. EuroHPC JU



Superkomputer Vega, Słowenia. Fot. Atos

EuroHPC PL – polski element EuroHPC JU



Na początku 2021 roku rozpoczęła się realizacja projektu „Narodowa Infrastruktura Superkomputerowa dla EuroHPC – EuroHPC PL”, którego celem jest zapewnienie nowoczesnej infrastruktury, w szczególności zasobów superkomputerów, specjalizowanych akceleratorów i usług laboratoriów na potrzeby środowiska akademickiego, podmiotów gospodarki i społeczeństwa w Polsce.

Projekt jest realizowany przez Konsorcjum w składzie:

- Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie – Akademickie Centrum Komputerowe Cyfronet – Lider Konsorcjum,
- Instytut Chemii Bioorganicznej Polskiej Akademii Nauk – Poznańskie Centrum Superkomputerowo-Sieciowe,
- Politechnika Gdańska – Centrum Informatyczne Trójmiejskiej Akademickiej Sieci Komputerowej,
- Politechnika Wrocławska – Wrocławskie Centrum Sieciowo-Superkomputerowe,
- Narodowe Centrum Badań Jądrowych,
- Instytut Informatyki Teoretycznej i Stosowanej Polskiej Akademii Nauk,
- Centrum Fizyki Teoretycznej Polskiej Akademii Nauk.

EuroHPC PL ma ułatwić i poszerzyć dostęp do wielkoskalowego przetwarzania danych zarówno dla naukowców, jak i dla przedsiębiorców i w ten sposób zwiększyć konkurencyjność polskich podmiotów na tle Europy i świata. Podstawowym elementem budowanej infrastruktury jest moc obliczeniowa dostarczana przez superkomputery. Ich zastosowanie jest składnikiem niezbędnym do prowadzenia badań naukowych w wielu dziedzinach i znacząco przyspiesza oraz obniża koszty procesów badawczych. Jest to szczególnie widoczne na początkowym etapie procesu badawczego, w którym zbiór możliwych do rozważenia kierunków badań jest tak duży, że ich klasyczne sprawdzenie (np. w laboratoriach) nie byłoby możliwe zarówno ze względu na czasochłonność jak i koszty. Zasoby superkomputerowe zostaną wsparte architekturą algorytmów kwantowych i akceleratorów neuromorficznych. Dzięki temu znacznie zwiększą się możliwości analizy i wizualizacji danych. Dodatkowo projekt zakłada szeroki zakres wsparcia użytkownika i szkolenia. Zadania w ramach projektu realizowane są w ramach czterech laboratoriów.

Głównym celem Laboratorium modelowania i równoległego przetwarzania danych w środowisku przedeksaskalowym jest zaprojektowanie i wdrożenie zaawansowanych systemów obliczeniowych dużej skali. Laboratorium zastosowań obliczeń hybrydowych adresuje potrzeby związane z zastosowaniami i wdrożeniem do praktycznego użycia komputerów kwantowych. Z kolei zadania Laboratorium zastosowania superkomputerów w medycynie skupiają się nad praktycznym zastosowaniem dużej mocy obliczeniowych w aplikacjach medycznych. Laboratorium efektywności energetycznej i obliczeniowej oprogramowania HPC pracuje nad rozwojem technik mających na celu redukcję kosztów operacyjnych maszyny, takich jak zużycie energii i zapotrzebowanie na chłodzenie, a także nad optymalizacją oprogramowania działającego w środowisku superkomputerowym.

Więcej informacji na temat EuroHPC PL znajduje się na stronie: <https://eurohpc.pl>.

LUMI

Konsorcjum LUMI

Polska, reprezentowana przez ACK Cyfronet AGH i MNiSW, dołączyła do kierowanego przez Finlandię konsorcjum LUMI, które opracowało jedną z trzech planowanych maszyn przed-eksaflopowych. Konsorcjum tworzą również: Belgia, Czechy, Dania, Estonia, Islandia, Norwegia, Szwecja i Szwajcaria.

LUMI, będące akronimem od The Large Unified Modern Infrastructure, jest również słowem z języka fińskiego oznaczającym „śnieg”. I podobnie jak śnieg, może całkowicie zmienić krajobraz – w tym przypadku badań opartych na danych. Połączenie unikalnej wiedzy i doświadczenia w budowie i obsłudze systemów superkomputerowych, zarządzaniu Big Data oraz tworzeniu i wykorzystywaniu zaawansowanego oprogramowania ma zaowocować stworzeniem łatwo dostępnego, wydajnego i bezpiecznego środowiska HPC o dużych możliwościach.

Komputer przed-eksaskalowy

Superkomputer LUMI został zainstalowany w centrum danych CSC w Kajaani w Finlandii. Konfiguracja LUMI została oparta na 64-rdzeniowych procesorach AMD EPYC™ oraz na procesorach graficznych (GPU) AMD Instinct™. Superkomputer ma udostępniać europejskiej nauce i gospodarce ponad 550 PFlops szczytowej mocy obliczeniowej, co odpowiada mocy 1,5 miliona nowoczesnych laptopów o dobrej konfiguracji, oraz 117 PB zasobów dyskowych, specjalizowane partycje i pakiety oprogramowania naukowego. Dzięki wykonywaniu większej liczby obliczeń w krótszym czasie i przetwarzaniu modeli bardziej precyzyjnych niż kiedykolwiek wcześniej, maszyna da naukowcom możliwość szybszego wykonywania znacznie bardziej skomplikowanych obliczeń i przetwarzania dużych zbiorów danych. Spodziewanym wynikiem będzie dynamiczny rozwój w wielu kierunkach badawczych – w zakresie m.in. medycyny spersonalizowanej, chemii, biologii, nanotechnologii, inżynierii materiałowej i innych dziedzin istotnych dla współczesnej gospodarki. W perspektywie najbliższych kilku lat LUMI ma stać się też jednym z najbardziej rozpoznawalnych narzędzi badawczych w zakresie AI – szczególnie *deep learning*.

“Superkomputer będzie można wykorzystać do rozwiązywania tych problemów nauki, które przy obecnie dostępnych maszynach pozostają nierozwiązane – mówi Marek Magryś, Zastępca Dyrektora Cyfronetu ds. Komputerów Dużej Mocy. – Takim flagowym przykładem są symulacje mózgu”.

Superkomputer LUMI ma wielkość porównywalną do kortu tenisowego i jest w 100% zasilany energią ze źródeł odnawialnych. Ciepło powstające jako efekt uboczny działania maszyny jest wykorzystywane do ogrzewania pobliskich budynków.

W momencie pełnego uruchomienia superkomputer LUMI będzie jednym z najszybszych systemów komputerowych na świecie.



Zdjęcie pierwszej fazy instalacji LUMI. Fot. Juha Torvinen, CSC

Dostęp do LUMI dla polskich naukowców

ACK Cyfronet AGH koordynuje prace związane z udostępnieniem polskim naukowcom zasobów oferowanych przez LUMI. Celem jest zapewnienie wydajnego i przyjaznego dla użytkownika środowiska pracy z superkomputerem, na bazie wypracowanych dobrych praktyk. Użytkownicy infrastruktury PLGrid mogą wnioskować o granty obliczeniowe na superkomputerze LUMI, podobnie jak przedstawiciele innych państw konsorcjum.

Konkurs na projekty pilotażowe

Pod koniec 2021 roku uruchomione zostaną pilotażowe granty obliczeniowe, które mają przetestować możliwości budowanej infrastruktury. Polska, jako członek konsorcjum, zgłosiła projekty o potencjalnie dużym znaczeniu, wymagające użycia dużej ilości rdzeni obliczeniowych oraz potężnych zasobów dyskowych. Projekty pilotażowe są wyłaniane na drodze konkursów i będą realizowane w dwóch etapach:

1. Projekty z obliczeniami opartymi o zasoby CPU (jesień 2021) – 2 projekty z Polski (*konkurs rozstrzygnięty*).
2. Projekty z obliczeniami opartymi o zasoby GPU (początek 2022) – 2 projekty z Polski.

Dostęp jest realizowany za pośrednictwem Portalu PLGrid (<https://portal.plgrid.pl>), a dodatkowe informacje znajdują się na stronie <https://cyfronet.pl/lumi>.

Więcej o LUMI: <https://www.lumi-supercomputer.eu/>.

Cyfronet w przedsięwzięciach wpisanych na Polską Mapę Infrastruktury Badawczej



Wśród strategicznych infrastruktur wpisanych w styczniu 2020 r. na Polską Mapę Infrastruktury Badawczej znajdują się dwa przedsięwzięcia zgłoszone przez ACK Cyfronet AGH jako inicjatora i koordynatora konsorcjum PLGrid: *Narodowa Infrastruktura Superkomputerowa dla EuroHPC* oraz *Narodowa Infrastruktura Chmurowa PLGrid dla EOSC*.*

Celem programu **Narodowa Infrastruktura Superkomputerowa dla EuroHPC** jest budowa infrastruktury obliczeniowej dla badań naukowych dotyczących rozwiązań odpowiadających na aktualne i przyszłe potrzeby polskiego społeczeństwa, środowiska naukowego i gospodarki.

Infrastruktura będzie opierać się o nowoczesne systemy superkomputerowe umożliwiające realizację zarówno tradycyjnych zadań symulacyjnych, jak i analizę danych z wykorzystaniem metod sztucznej inteligencji. Produkcyjne systemy obliczeniowe budowane w ramach projektu będą należeć do czołówki światowych superkomputerów. Poza infrastrukturą obliczeniową projekt zapewni też dostęp do specjalistycznych szkoleń oraz eksperckiego wsparcia technicznego dla użytkowników z nauki i gospodarki, a także niezbędne procedury alokacji i rozliczania wykorzystania zasobów.

Projekt wpisuje się bezpośrednio w ramy międzynarodowej inicjatywy EuroHPC – European High-Performance Computing opublikowanej w komunikacie Komisji Europejskiej nr COM/2018/08 final – 2018/03 (NLE). EuroHPC jest przedsięwzięciem mającym na celu utworzenie europejskiego systemu komputerów dużej mocy klasy eksaskalowej, unikatowych na skalę światową, zbudowanych w oparciu o technologie rozwijane w Europie. Przedsięwzięcie będzie realizowane przez Konsorcjum PLGrid.

W efekcie realizacji przedsięwzięcia Narodowa Infrastruktura Superkomputerowa dla EuroHPC powstanie infrastruktura do prowadzenia badań na potrzeby nauki, gospodarki i społeczeństwa czerpiąca z najnowocześniejszych technologii w zakresie HPC wypracowanych w ramach międzynarodowej kooperacji EuroHPC. Infrastruktura będzie oferowała usługi w zakresie masywnie równoległych symulacji komputerowych, wysoce wydajnego przetwarzania zbiorów danych, wykorzystania metod sztucznej inteligencji, oprogramowania i narzędzi wysokiej produktywności, w tym wizualizacji danych, oraz wsparcia i szkolenia użytkowników. Poza głównymi systemami obliczeniowymi w skład infrastruktury wejdą także mniejsze systemy testowe i badawcze, służące weryfikacji nowych technologii procesorowych, akceleratorowych, pamięciowych i sieciowych w kontekście wykorzystania ich do budowy systemów produkcyjnych, a także prowadzeniu prac badawczych i rozwojowych w zakresie efektywnych infrastruktur HPC.

Usługi będą oferowane za pośrednictwem infrastruktury PLGrid, która integruje większość dostępnych w kraju zasobów obliczeniowych, co ułatwi proces alokacji zasobów i wsparcia użytkowników. Zrealizowana zostanie także integracja z europejskimi systemami preeskalowymi i petaskalowymi udostępnionymi w ramach programu Euro-HPC, w szczególności z maszyną budowaną przez konsorcjum LUMI, którego Polska jest członkiem.

Narodowa Infrastruktura Chmurowa PLGrid dla EOSC to program wykorzystania zasobów chmurowych do badań naukowych odpowiadających na aktualne i przyszłe potrzeby polskiego społeczeństwa, środowiska naukowego i gospodarki. Zakres tych badań obejmuje m.in. dane, infrastruktury i platformy przetwarzania danych oraz efektywne algorytmy i dedykowane aplikacje.

Podstawą programu są wymagania społeczeństwa, gospodarki i polskich badaczy, w szczególności współpracujących w ramach międzynarodowych zespołów naukowych. Grupy te wymagają zaawansowanych środowisk integracji rozproszonych zasobów: oprogramowania, infrastruktury i dedykowanych usług. Wymagania te mogą być spełnione jedynie przez zaawansowane technologie IT, połączone z zasobami obliczeniowymi, pamięci masowych i danych. Współdziałanie tych wszystkich elementów w ramach elastycznego ekosystemu umożliwiają technologie chmurowe.

Narodowa Infrastruktura Chmurowa PLGrid dla EOSC jest częścią ekosystemu Europejskiej Chmury dla Otwartej Nauki (EOSC, Deklaracja z dnia 26.10.2017). Polska rozwija obecnie dwa kluczowe komponenty tego federacyjnego, globalnie dostępnego i multidyscyplinarnego środowiska: Onedata – system jednolitego udostępniania i zarządzania danymi oraz Portal EOSC. Jako część europejskiego ekosystemu Narodowa Infrastruktura Chmurowa PLGrid będzie oferować zaufane i otwarte środowiska dla użytkowników w całym cyklu życia danych. Pozwoli to na publikację, wyszukiwanie, wykorzystywanie i wielokrotne użycie zgromadzonych danych, narzędzi, oprogramowania i innych wyników prac przez naukowców, gospodarkę i całe społeczeństwo.

Badania zaplanowane w ramach Narodowej Infrastruktury Chmurowej PLGrid dla EOSC pozwolą na opracowanie, walidację i w konsekwencji udostępnienie usług (ogólnych i dedykowanych), a tym samym wykorzystanie nowoczesnych technologii i efektywnych technik zarządzania, przetwarzania i ponownego wykorzystania danych przez środowiska naukowe, podmioty gospodarki i społeczeństwo.

W tym celu Infrastruktura będzie dostarczać:

- technologie rozwiązań dla środowisk rozproszonych, w tym chmurowych, obejmujące zarządzanie usługami w środowisku rozproszonym, automatyzację złożonych procesów, integrację platform i infrastruktury badawczych,
- oparte o krajowe technologie bezpieczne udostępnianie danych i zarządzanie nimi, zgodne ze standardami EOSC,
- badania specyficznych potrzeb użytkowników w zakresie wielkoskalowego przetwarzania danych w środowisku rozproszonym, w tym: przetwarzanie „blisko danych”, wykorzystanie paradygmatu „data lakes” wraz z nowymi modelami analizy, skalowalne zasoby w środowisku rozproszonym,
- weryfikację rozwiązań przygotowywanych na potrzeby społeczeństwa, nauki i gospodarki w zaawansowanych aplikacjach,
- katalog ogólnych serwisów i usług chmurowych dla gospodarki, nauki i społeczeństwa będących wynikiem prac badawczych i rozwojowych we współpracy z EOSC.

Pierwszy etap tych prac jest obecnie realizowany w ramach zasobów infrastruktury PLGrid.

**Materiały z broszury MNiSW pt. „Polska Mapa Infrastruktury Badawczej”.*



W ramach fazy operacyjnej infrastruktury EPOS udostępniane są operacyjnie zasoby węzłów tematycznych. Polska udostępnia węzeł tematyczny hazardu antropogenicznego, tzw. TCS AH (ang. Thematic Core Services Anthropogenic Hazard). Głównym zadaniem Cyfronetu w projekcie jest operacyjne udostępnianie portalu i zintegrowanych z nim zasobów TCS AH (danych, oprogramowania i mocy obliczeniowych do ich analizy).



Głównym celem projektu EuroCC jest stworzenie europejskiego systemu wsparcia technologicznego opartego o ściśle powiązane ze sobą Narodowe Centra Kompetencji w poszczególnych krajach europejskich, co pozwoli na odnoszenie korzyści środowisku akademickiemu, sektorowi przedsiębiorstw (zwłaszcza MŚP) oraz administracji publicznej z dostępnej wiedzy specjalistycznej, doświadczenia i zasobów EuroHPC.



W ramach projektu PROTEUS-RS przewiduje się wykonanie modeli oraz szeregu symulacji numerycznych w oparciu o architektury komputerowe wysokiej wydajności na potrzeby projektowania procesu wytwarzania metalowych elementów długich.



Celem projektu EPOS-SP jest opracowanie i wdrożenie założeń pozwalających na zapewnienie trwałości infrastruktury EPOS wytworzonej w poprzednich projektach (EPOS-PP oraz EPOS-IP).



Projekt EOSC Enhance miał na celu zbudowanie ulepszonej, bardziej zintegrowanej wersji Portalu EOSC, co umożliwiło poprawę rozszerzenia rozwiązań ułatwiających odszukiwanie europejskich usług naukowych oraz zbiorów danych otwartej nauki.



Projekt EOSC-Synergy wprowadza standardy EOSC dla infrastruktur krajowych w dziewięciu krajach Unii Europejskiej. Stanie się to poprzez zharmonizowanie polityk oraz rozszerzenie dostępu do infrastruktur badawczych, danych naukowych i serwisów dziedzinowych.



Celem projektu jest utworzenie w Krakowie ośrodka medycyny obliczeniowej. Centrum będzie głównym motorem europejskiego postępu w tym szybko rozwijającym się sektorze, opracowując zaawansowane metody inżynierskie do zapobiegania, diagnozowania i leczenia chorób oraz zaspokajając ogólnoswiatową potrzebę radykalnie usprawnionych systemów opieki zdrowotnej.



Głównym zadaniem projektu PRACE-6IP jest implementacja nowych rozwiązań oraz utrzymanie operacyjności środowiska PRACE w obszarze europejskich infrastruktur obliczeniowych HPC.



W ramach projektu stworzony zostanie system wspierania decyzji klinicznych (Clinical Decision Support System – CDSS) dla leczenia nowotworów (nauroblastoma, glejak) występujących u dzieci.



Celem projektu było przygotowanie uruchomienia produkcyjnej infrastruktury dla otwartej nauki w Europie i praktyczne zastosowanie rozwiązań wypracowanych w ramach projektu EOSC-Pilot do realnego środowiska wielkiej skali rozproszonego na większość krajów europejskich.



Rozwiązania opracowane w projekcie są przełomowym krokiem na drodze do stworzenia innowacyjnych serwisów przetwarzania eksaskalowych danych, maksymalizujących korzyści, jakie niosą nowoczesne systemy obróbki danych.



Celem projektu EGI-Engage była popularyzacja idei „Open Science Commons”, czyli paradygmatu otwartej nauki dla wzbogacenia doświadczenia, wiedzy i współpracy pomiędzy grupami naukowców w Europie. Implementacja paradygmatu oparta została o sfederowane w ramach EGI infrastruktury obliczeniowe, danych, komunikacji, a także o wiedzę i doświadczenie społeczności naukowych je wykorzystujących.



W ramach projektu tworzone i rozwijano nowe gridowe narzędzia i usługi służące do uruchamiania interaktywnych aplikacji, wymagających dużej mocy obliczeniowych do przeprowadzania wielkoskalowych obliczeń rozproszonych.



Zadaniem projektu jest opracowanie i dostarczenie produkcyjnych usług przechowywania, dostępu oraz zabezpieczania danych i zarządzania metadanymi, a także integracji rozwiązań dla przetwarzania dużych i złożonych wolumenów danych na bazie rozproszonej e-infrastruktury.



Celem projektu jest budowa specjalistycznej infrastruktury ogólnego przeznaczenia na potrzeby obliczeń wielkoskalowych umożliwiającej podejmowanie wyzwań badawczych w kluczowych z punktu widzenia polskiego społeczeństwa, środowiska naukowego i gospodarki obszarach. Projekt jest polskim etapem rozwoju programu EuroHPC.



Projekt zakłada budowę unikalnych laboratoriów badawczych w oparciu o krajową sieć światłowodową PIONIER. Głównym celem projektu jest zbudowanie i udostępnienie platform dla jednostek naukowych, przedsiębiorców i innych podmiotów zainteresowanych prowadzeniem badań naukowych oraz prac rozwojowych w oparciu o nową, ogólnopolską infrastrukturę badawczą.



Celem bezpośrednim projektu jest budowa specjalistycznej e-infrastruktury przetwarzania danych, umożliwiającej optymalne wykorzystanie usług specjalistycznych oraz nowej generacji służących do stymulowania nowych obszarów zastosowań w nauce, gospodarce, edukacji oraz życiu społecznym.



W ramach projektu zostaną zwiększone funkcjonalności infrastruktury badawczej EPOS-PL. Powstanie nowe Centrum Infrastruktury Badawczej, nowy poligon pomiarowy oraz WNiP: „Platforma IT do badań metodami Sztucznej Inteligencji (EPOS-AI)”.



Zadaniem projektu jest wytworzenie usług infrastruktury obliczeniowej oraz składowania danych na potrzeby projektu PRACE w ramach 6-ciu dedykowanych laboratoriów: 1) L. przetwarzania HPC oraz chmurowego, 2) L. dostępu do infrastruktury przetwarzania, 3) L. zarządzania i monitorowania usług, 4) L. usług zarządzania danymi, 5) L. zarządzania rozproszonymi danymi oraz transparentnego dostępu do danych, 6) L. bezpieczeństwa infrastruktury.



Głównym zadaniem projektu było udostępnienie pochodzących z satelitów Sentinel sieci Copernicus danych satelitarnych. W ramach projektu powstała infrastruktura do automatycznego pobierania danych bezpośrednio z satelitów, ich bezpiecznego przechowywania oraz udostępniania na potrzeby nauki, administracji i szkoleń.



Zadaniem projektu była budowa krajowej infrastruktury badawczej z zakresu nauk o Ziemi oraz jej integracja z międzynarodowymi bazami danych, serwisami i usługami realizowanymi w ramach europejskiego programu EPOS (European Plate Observing System).



Celem projektu był rozwój wyspecjalizowanego centrum kompetencji technologicznych w zakresie rozproszonych infrastruktur obliczeniowych, ze szczególnym uwzględnieniem technologii gridowych, obliczeń w chmurze oraz infrastruktur wspierających obliczenia na wielkich zbiorach danych. W rezultacie, użytkownikom zostały zaoferowane wielkie moce obliczeniowe i duże zasoby pamięci. Infrastruktura PLGrid została wzbogacona o nowe usługi bazowe i usługi dla użytkowników końcowych.



W trakcie projektu miała miejsce integracja wybranych usług dostępnych w sieci PIONIER i opracowano nowe usługi, w zakresie zwiększania niezawodności i bezpieczeństwa.



Celem projektu była rozbudowa 21 środowiskowych sieci teleinformatycznych nauki dla zapewnienia instytucjom naukowym dostępu do nowoczesnej oraz bezpiecznej infrastruktury sieciowej i umożliwiającej łączność z jednostkami naukowymi całego świata, wspierając polskie grupy naukowców.



Głównym zadaniem projektu było stworzenie i uruchomienie pięciu usług działających w oparciu o sieć PIONIER: wideokonferencji, eduroam, usług kampusowych, powszechnej archiwizacji oraz naukowej interaktywnej telewizji HD.



Młodzi naukowcy w Cyfronecie

Organizowany corocznie konkurs na najlepszą pracę doktorską zrealizowaną w oparciu o zasoby obliczeniowe ACK Cyfronet AGH jest jedną z ważniejszych tradycji w naszym Centrum. Ocenie konkursowej poddawana jest wartość naukowa zgłoszonej pracy doktorskiej, możliwość jej praktycznego zastosowania oraz zakres wykorzystania zasobów obliczeniowych i pamięci składowania danych w Cyfronecie. Na przestrzeni ostatnich lat Konkurs stał się ważnym narzędziem promującym badania prowadzone przez młodych naukowców, a do kolejnych edycji Konkursu został zgłoszony szereg prac opisujących ciekawe zagadnienia naukowe z takich dziedzin jak chemia, fizyka czy energetyka. Młodzi badacze łączyli w swoich badaniach wyniki bezpośrednich pomiarów w laboratorium i obliczenia na komputerach dużej mocy. W swojej pracy korzystali z różnych narzędzi uruchamianych na szerokim wachlarzu architektur obliczeniowych oferowanych przez Cyfronet.

Prezentacja najważniejszych tez prac doktorskich Laureatów Konkursu miała miejsce w trakcie Dnia Otwartego Cyfronetu w listopadzie 2021. Z przyjemnością prezentujemy fragmenty kilku wybranych wywiadów z uczestnikami Konkursu, które znajdziecie Państwo na kolejnych stronach.

Zapraszamy serdecznie do uczestnictwa w kolejnej edycji Konkursu!

<http://www.cyfronet.pl/konkurs>



Laureaci jednej z poprzednich edycji



dr Kévin Almeida Cheminant

Rozmowa z autorem pracy:

„Search for ultra-high energy photons through preshower effect with gamma-ray telescopes”

Co zdecydowało o tym, że poświęcił Pan rozprawę doktorską zagadnieniu fotonów gamma o skrajnie wysokich energiach?

Promienie kosmiczne o ultrawysokiej energii (UHECR) to naładowane cząstki przemieszczające się przez Wszechświat o energii porównywalnej z energią profesjonalnie zaserwowanej piłki tenisowej. Taka energia jest poza zasięgiem obecnych możliwości akceleratorów cząstek elementarnych, a zatem UHECR są bardzo wartościowe, gdyż umożliwiają nam zrozumienie fizyki ekstremalnie wysokich energii. Niemniej, ich pochodzenie pozostaje owiane tajemnicą. Jednak w połączeniu z UHECR są często emitowane również fotony o ultrawysokiej energii. Poszukiwanie tych fotonów ma kilka zalet. Jako cząstki neutralne nie są odchylane przez pola magnetyczne i odzyskanie ich źródła emisji rzuciłoby trochę światła na pochodzenie UHECR. Po drugie, różne scenariusze produkcji UHECR prowadzą do różnych oczekiwań odnośnie strumienia fotonów UHE. Niestety, żaden foton UHE nigdy nie został zaobserwowany. Moja praca doktorska była poświęcona opracowaniu nowej metody wyszukiwania tych fotonów za pomocą teleskopów promieniowania gamma, które zostały zaprojektowane do obserwacji fotonów GeV-TeV. Pokazaliśmy, że te teleskopy mogą skutecznie wykrywać fotony UHE, dzięki czemu do poszukiwań już zainicjowanych przez Obserwatorium Pierre Auger i Telescope Array dołączyła grupa teleskopów gamma rozsianych po całym świecie.

Od dłuższego czasu jest Pan uczestnikiem projektu CREDO. Jak przybliżyłby go Pan osobie niezwiązanej z dziedziną?

CREDO to skrót od Cosmic-Ray Extremely Distributed Observatory. Idea, która stoi za tym projektem, polega na stworzeniu ogólnosiwiatowej sieci detektorów promieniowania kosmicznego w celu poszukiwania możliwych korelacji między odległymi detekcjami tego promieniowania. Aby stworzyć taką sieć, należy wziąć pod uwagę każdy potencjalny detektor promieniowania kosmicznego, zarówno profesjonalny (detekcja neutronów, ciemnej materii, eksperymenty z promieniowaniem gamma, itp.) jak i aparaturę edukacyjną (CosmicWatch, itp.). W tym celu, w ramach współpracy CREDO opracowano aplikację na smartfony, która może zmienić ich czujniki kamery w detektor cząstek. Zgodnie z filozofią CREDO, wszystkie dane i analizy są w pełni dostępne dla każdego uczestnika projektu i dzięki temu umożliwiają dostrzeżenie potencjalnego związku między promieniami kosmicznymi a innymi dziedzinami fizyki, takimi jak np. biofizyka (wpływ promieniowania na organizm żywy) lub geofizyka (związek między promieniowaniem i trzęsieniami ziemi). Więcej informacji o CREDO można znaleźć na stronie <https://credo.science>.

W pracy wykorzystuje Pan zarówno dane pochodzące z obserwacji jak i te powstałe w wyniku symulacji komputerowych. Jak ocenia Pan zaawansowanie i przydatność modelowania komputerowego w astrofizyce i jak zasoby Cyfronetu wpisują się w oczekiwania astrofizyków?

Promieniowanie kosmiczne oraz gamma, oddziałujące z atmosferą ziemską wytwarzają kaskady zawierające setki tysięcy cząstek wtórnych, które są wynikiem wielu różnych rodzajów interakcji. Ponieważ modele analityczne nie dają pełnego opisu tych kaskad, musimy polegać na rozbudowanych symulacjach Monte-Carlo, które wymagają dużych zasobów obliczeniowych. Symulacje te mogą opisać zachowanie każdej cząstki, a także uwzględniać probabilistyczny charakter oddziaływań cząstek. Dzięki nim możemy przyjąć pewne założenia i testować różne modele fizyczne, które następnie są porównywane z zebranymi danymi. W mojej pracy zasoby Cyfronetu pozwoliły mi na uruchamianie bardzo dużej liczby niezwykle wymagających obliczeniowo symulacji Monte-Carlo. Bez nich nie uzyskałbym wystarczającej liczby statystyk, aby zbudować moje

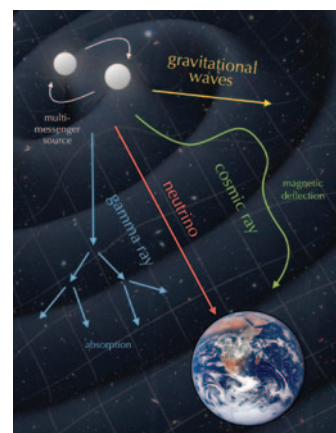
modele w rozsądnym czasie. Co więcej, w miarę jak coraz lepiej rozumiemy interakcje fizyczne i dążymy do bardziej szczegółowych opisów zjawisk fizycznych, centra superkomputerowe będą miały coraz większą rolę do odegrania w nadchodzących latach. Jestem przekonany, że wiedza i doświadczenie kadry Cyfronetu będą niezwykle cenne dla każdego naukowca potrzebującego dużych zasobów komputerowych.

W oparciu o Pańskie doświadczenie i wiedzę, a także wyniki zawarte w rozprawie, jakie są Pańskie przewidywania dotyczące rozwoju dyscypliny?

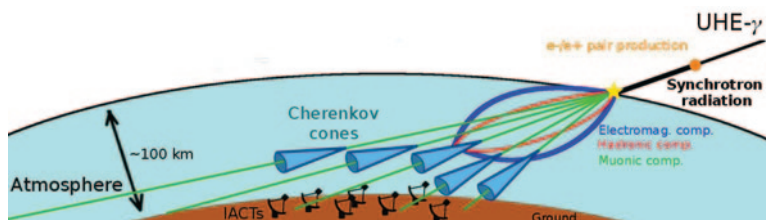
Chociaż istniejące eksperymenty fizyki astrocząstek przyniosły bezprecedensowy przyrost wiedzy o najwyższych obserwowalnych energiach we Wszechświecie, wciąż dążymy do bardziej czułych detektorów o lepszej rozdzielczości kątowej i energetycznej. Najnowsza generacja teleskopów gamma wykorzystywana w ramach przedsięwzięcia Cherenkov Telescope Array (CTA) zapewni lepsze zrozumienie znanych źródeł promieniowania gamma, ale także umożliwi wyszukiwanie UHECR. W fizyce promieniowania kosmicznego o ultra wysokiej energii dodatkowe detektory dostarczą więcej informacji o interakcjach cząstek w pękach i większej ilości danych statystycznych, potrzebnych do zrozumienia natury tego promieniowania. Korzystamy zatem również z eksperymentów, takich jak IceCUBE, KM3Net lub Baikal-GVD, gdyż dają one nowy wgląd we Wszechświat ekstremalnych energii. Przyszłość dyscypliny nieuchronnie leży w nieustannej komunikacji i wymianie danych, a społeczność fizyków cząstek i badaczy już zapoczątkowała ten proces. Dzięki systemowi alertów rozblyskujące źródła astrofizyczne są teraz obserwowane w ramach różnego rodzaju eksperymentów, co pozwala nam zdobyć bardzo zróżnicowane i ważne dane na temat zjawisk fizycznych zachodzących w tych źródłach.

Co chciałby Pan przekazać osobom rozważającym dopiero podjęcie studiów doktoranckich?

Myślę, że najlepszą radą, jaką mogę dać każdemu rozpoczynającemu studia doktoranckie, to nie bać się poznawania nowych metod lub nowych tematów. Bycie nowicjuszem w wyspecjalizowanych dziedzinach może być przerażające, a w moim przypadku często czułem, że wszystko, co wymyślę, będzie po prostu mało interesujące. Otóż nic bardziej mylnego! Myślę, że każdy projekt badawczy wymaga odrobiny kreatywności oraz ciekawości i nie należy obawiać się eksperymentowania z własnymi pomysłami, nawet jeśli finalnie będziemy musieli zmienić koncepcję. Zadawanie pytań i ciągłe bycie na bieżąco z dziedziną mogą również sprawić, że studia doktoranckie będą bardzo satysfakcjonującym doświadczeniem. Wreszcie, rozmowy i uczestnictwo w prezentacjach to doskonałe ćwiczenie i dobry sposób na efektywne podsumowanie swoich własnych badań, ale także pozyskiwanie cennych uwag i komentarzy od ekspertów.



Obiekty astrofizyczne mogą przysyłać różnego rodzaju informacje, w tym neutrino, fale grawitacyjne, promienie kosmiczne i promienie gamma. Obserwacja tych obiektów za pomocą różnego rodzaju eksperymentów w celu wykrycia tych informacji zapewnia nam lepsze zrozumienie zjawisk fizycznych zachodzących przy najwyższych energiach.



Foton o ultrawysokiej energii oddziałujący z polem geomagnetycznym wytwarza parę elektron/pozyton, która emituje fotony wtórne (promieniowanie synchrotronowe). Ten zbiór cząstek, zwany także deszczem wstępnym, wytwarza kaskadę atmosferyczną zawierającą setki tysięcy cząstek. Taką kaskadę można wykryć na ziemi za pomocą teleskopów IACT (Imaging Atmospheric Cherenkov Telescopes) poprzez niebieskawe promieniowanie Czerenkowa emitowane w powietrzu przez naładowane cząstki (elektrony, miony, itp.).



dr Klemens Noga

Rozmowa z autorem pracy:

„Badanie oddziaływań elektronowych w związkach molibdenu o mieszanej walencyjności metodami chemii kwantowej”

Jaki aspekt chemii obliczeniowej jest dla Pana najbardziej interesujący?

Najciekawszym dla mnie aspektem chemii obliczeniowej jest to, że przeprowadza się badania na różnych modelach, które stosunkowo łatwo można modyfikować. Dzięki temu możemy analizować wpływ różnych czynników na własności układów chemicznych. Dodatkowo, dzięki coraz to nowocześniejszej technologii obliczeniowej można analizować coraz bardziej złożone modele.

Co zdecydowało o wyborze skorpionianowych kompleksów molibdenu jako przedmiotu badań?

Temat moich badań początkowo powstał z potrzeby chwili. Współpracownicy z Politechniki Krakowskiej pracowali nad zrozumieniem mechanizmów rządzących zachowaniem skorpionianowych kompleksów molibdenu, w szczególności zawierających więcej niż jedno centrum metaliczne. Do tego potrzebna była analiza czynników wpływających na strukturę elektronową, a tę można skutecznie modelować przy użyciu komputerów.

Same skorpioniany, z racji wszechstronności ligandów, które mogą je tworzyć, jak również sposobów, na które poszczególne centra metaliczne mogą być połączone, okazały się być bardzo ciekawym, a zarazem skomplikowanym przedmiotem badań. Dodatkowo, ich zdolność do tworzenia związków o mieszanej walencyjności, gdzie poszczególne centra metaliczne znajdują się na innym formalnym stopniu utlenienia, a niesparowany elektron może być mniej lub bardziej współdzielony pomiędzy centra tworzące taki kompleks, powoduje, że są dobrymi systemami do analizy wewnątrzcząsteczkowych transferów elektronu. W czasie eksperymentów doświadczalnicy zaobserwowali ciekawe zjawisko elektrokatalizy rozpadu chloroformu przez niektóre ze skorpionianów. Badanie również i tego procesu było bardzo ciekawym wyzwaniem.

Metodologia przeprowadzonych przeze mnie obliczeń, w tym przepisy na dokładne obliczanie potencjałów utleniająco-redukujących centrów metalicznych, jak i analiza wpływu pewnych oddziaływań zwanych oddziaływaniami dyspersyjnymi, mogą być i są stosowane dla innych klas układów.

Jakie możliwości w czasie badań teoretycznych dała Panu wspomniana współpraca z zespołem doświadczalnym?

Bliska współpraca z eksperymentatorami – zespołem Piotra Romańczyka i Stefana Kurka z Politechniki Krakowskiej – była w moim przypadku kluczowa. Dzięki temu mogliśmy wzajemnie uzupełniać nasze braki w zrozumieniu różnych aspektów rozwiązywanych problemów, a ja mogłem bezpośrednio konfrontować rezultaty moich eksperymentów obliczeniowych z danymi doświadczalnymi.

Przy badaniach korzystał Pan z kilku kolejnych superkomputerów Cyfronetu. W jaki sposób udostępniane zasoby wpłynęły na Pańskie badania?

Po prostu bez zasobów Cyfronetu nie byłbym w stanie przeprowadzić moich badań. Każda kolejna generacja superkomputerów dawała nowe możliwości obliczeniowe, dzięki którym mogłem badać bardziej złożone układy lub wykonywać obliczenia

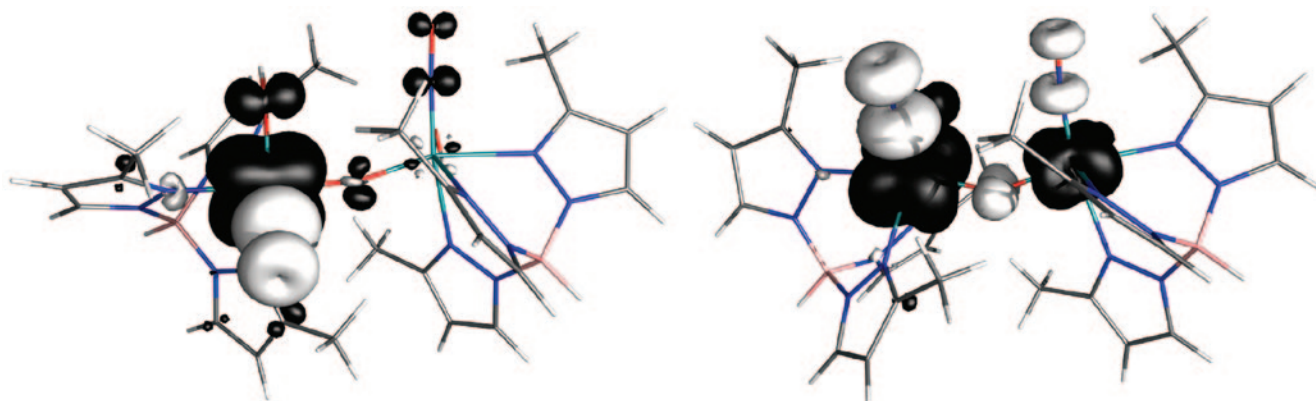
przy użyciu metod umożliwiających uzyskanie dokładniejszych wyników, bardziej zgodnych z dostępnymi danymi doświadczalnymi.

Czy na postawie procesu własnych badań pokusiłby się Pan o prognozę, w jakich kierunkach prowadzą naukę współczesne osiągnięcia technologiczne i odkrycia teoretyczne?

W ostatnich latach można zauważyć zwiększone znaczenie badań *in silico* w chemii, biochemii oraz naukach pokrewnych. Dzięki wykorzystywaniu superkomputerów można szybciej i bardziej efektywnie badać ciekawe układy, m.in. potencjalne leki czy też układy katalityczne. Również tak zwane uczenie maszynowe przyczyniło się do otrzymania interesujących wyników naukowych w naukach biochemicznych, umożliwiając np. przewidywanie struktur przestrzennych białek. Myślę, że będziemy mogli tworzyć lepsze, bardziej złożone modele interesujących nas układów, a to da nam lepszy wgląd w najważniejsze procesy chemiczne i biochemiczne.

Badania prowadzące do rozprawy doktorskiej są rozłożone w czasie i kończy się je z innym zestawem wiedzy i umiejętności, niż ten, z którym się je zaczynało. Jakiej porady udzieliłby Pan młodym naukowcom na początku tej drogi?

Na pewno poleciłbym, by pierwszy okres studiów poświęcili na rozbudowywanie swojej wiedzy o dostępnych metodach i technikach badawczych oraz nawiązali jak najszersze kontakty z różnymi naukowcami. Dzięki temu już na etapie doktoratu można uczestniczyć w bardzo ciekawych projektach badawczych, zobaczyć sposoby pracy w wielu grupach badawczych oraz mieć możliwość obserwacji i współuczestnictwa w opracowywaniu różnych raportów i publikacji. Dopiero w dalszej kolejności, na podstawie wcześniejszych doświadczeń, wybrałbym szczegółową tezę pracy, nad którą należy skoncentrować się w kolejnych latach. Na pewno każdemu rozpoczynającemu taką przygodę poradziłbym, by jak najwcześniej starał doskonalić się w opisywaniu swoich badań i przygotowaniu publikacji naukowych. Ta umiejętność przyda się nie tylko w trakcie pisania pracy doktorskiej.



Kontur gęstości spinowej uzyskanej dla jednego z badanych skorpionianów molibdenu w geometriach minimum energetycznego (po lewej) oraz maksymalnej delokalizacji niesparowanego elektronu (po prawej)



dr inż. Tomasz Prokop

Rozmowa z autorem pracy:

„Three-dimensional numerical analysis of transport phenomena in a Positive Electrolyte-Negative assembly of a Solid Oxide Fuel Cell”

Co zdecydowało o tym, że poświęcił Pan rozprawę doktorską zagadnieniu pozyskania energii elektrycznej z ogniw SOFC?

Do podjęcia badań nad ogniwami z elektrolitem stałotlenkowym (ogniw typu SOFC) zachęcili mnie jeszcze w czasie studiów magisterskich moi promotorzy. Jest to temat pełen ciekawych, wielodyscyplinarnych problemów badawczych. Co więcej, użycie ogniw paliwowych typu SOFC to jeden z najwydajniejszych sposobów wykorzystania energii paliw kopalnych. Wyróżniają się szczególnie wysoką sprawnością, a oparte o nie urządzenia mogą być zasilane szerokim zakresem paliw, w tym sieciowym gazem ziemnym. Oparte o SOFC pomocnicze źródła zasilania o mocy 1-250 kW (stosowane przy budynkach mieszkalnych, biurach, bądź serwerowniach) dorównują sprawnością kilkusetmegawatowym blokom zawodowych elektrowni gazowych. Są przy tym ciche, nie emitują trujących zanieczyszczeń, a ciepło odpadowe, dzięki wysokiej temperaturze pracy, może być wykorzystane do celów grzewczych. Proponowane są również oparte o SOFC elektrownie zawodowe, które w porównaniu do elektrowni konwencjonalnych, zużywałyby o ok. jedną trzecią mniej paliwa na jednostkę wyprodukowanej energii elektrycznej (i co za tym idzie, emitowały stosownie mniejszą ilość zanieczyszczeń).

Jak badania mogą się przełożyć na wdrożenia praktyczne?

Moje badania dotyczą modelowania zjawisk zachodzących w mikroskopijnej strukturze ogniw SOFC. Opracowany przeze mnie program do symulacji komputerowej ogniw może pomóc przy projektowaniu elektrod. Program wykorzystuje cyfrowe reprezentacje mikrostruktur elektrod zbadanych laboratoryjnie lub wygenerowanych komputerowo. Pozwala przeprowadzić analizę strat związanych m.in. z oporem przepływu prądu, czy nieefektywnością dostarczania reagentów gazowych. Symulacja pozwala przewidzieć, jak zmiana określonych cech ogniwa (parametrów materiałów, udziału poszczególnych składników, grubości warstw, itp.) wpłynie na wyżej wspomniane straty, bez konieczności produkcji prototypu i przeprowadzania żmudnych testów. Wierzę, że moje badania przyczynią się do zmniejszenia kosztów rozwoju ogniw SOFC i zwiększenia ich sprawności, a co za tym idzie – poprawy rentowności opartych o nie komercyjnych rozwiązań.

Jakie dalsze wyzwania stoją przed Panem i innymi naukowcami badającymi ogniwa paliwowe?

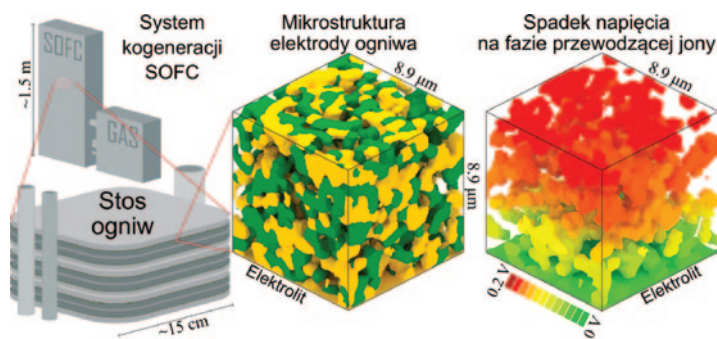
Główne wyzwanie przemysłu SOFC to wypracowanie technik masowej produkcji i zapewnienie jakości w trakcie pracy urządzenia. Ponieważ ogniwa SOFC pracują nawet w 1100°C, konieczne jest stosowanie materiałów odpornych na wysoką temperaturę i korozję oraz odpowiednich uszczelnień. Ogniwa SOFC nie są jeszcze zbyt trwałe, a występujące w nich podczas ogrzewania i chłodzenia naprężenia termiczne sprawiają, że można je włączyć i wygasić stosunkowo niewiele razy w trakcie życia urządzenia. Co więcej, mikrostruktura elektrod ogniw ewoluuje na przestrzeni tysięcy godzin funkcjonowania, co wpływa istotnie na parametry pracy całego systemu. Proces ten nie jest jeszcze w pełni poznany. Ważny dla szerszej implementacji ogniw paliwowych jest również rozwój technik związanych z oczyszczaniem, przetwarzaniem, magazynowaniem i dostarczaniem paliw.

Pańska praca doktorska ma charakter międzynarodowy, dosłownie międzykontynentalny. Czy może Pan rekomendować takie podejście Koleżankom i Kolegom rozpoczynającym przygodę ze studiami doktoranckimi?

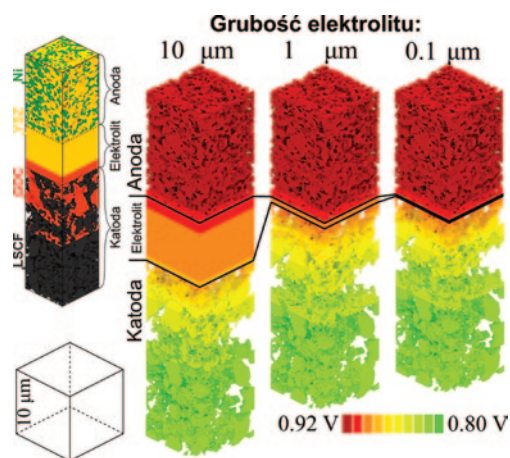
Zdecydowanie. Współpraca z innymi zespołami jest w nauce nieunikniona, a bardzo często ośrodek, który najlepiej odpowiada naszym zainteresowaniom, znajduje się w innym kraju. Powiedziałbym, że jest to norma, może poza dużymi krajami przeznaczającymi znaczące środki finansowe na badania. Podczas badań często koncentrujemy się na wąskim zagadnieniu – nierzadko danym problemem zajmuje się tylko kilka-kilkanaście zespołów na całym świecie. Warto nawiązać kontakty z tymi ludźmi – mogą znać sposoby na pokonanie przeszkód, z którymi się zmagamy, a nasze wyniki mogą być z kolei interesujące dla nich. Współpraca zagraniczna umożliwia dostęp do większej gamy staży naukowych, stypendiów, a także rzadkiej aparatury naukowej. Zwiększa widoczność naszych publikacji i pomaga określić nowe, interesujące kierunki badań.

Czy mógłby Pan na koniec krótko scharakteryzować japońskie podejście do kształcenia uniwersyteckiego i prowadzenia badań? Czy istnieją jakieś uderzające różnice kulturowe?

Japońscy magistranci nie uczestniczą zwykle w zorganizowanych zajęciach, pracują podobnie jak doktoranci. Są przypisani do grupy badawczej danego profesora, mają własne miejsca w biurze i dostęp do sprzętu laboratoryjnego. Podczas mojego stażu w Tokio średnio raz w tygodniu odbywały się seminaria, podczas których studenci, doktoranci i pracownicy omawiali swoje postępy oraz nowe wyniki badań opisane w literaturze. Spotkania te zajmowały kilka godzin, ale były bardzo interesujące. Dość często organizowane są wydarzenia integracyjne dla całego zespołu, np. wyjazdy turystyczne. Typową okazją jest przyjęcie lub pożegnanie członka zespołu. Ciekawe jest podejście do utrzymania czystości – personel sprzątający uczelni zajmuje się tylko korytarzami i innymi częściami wspólnymi, o biura i laboratoria dbają sami pracownicy i studenci. Zdziwić może również fakt, że Japończycy często drzemią w biurach i jest na to powszechne przyzwolenie. Drzemka interpretowana jest jako oznaka zmęczenia ciężką i sumienną pracą. Jest to szczególnie zrozumiałe na uczelni, gdzie pracowników i doktorantów można nierzadko spotkać nawet w środku nocy.



Mikrostruktura elektrody ogniwa paliwowego typu SOFC w kontekście przykładowego urządzenia. Każda faza mikrostruktury posiada właściwą sobie funkcję – wyszczególnione straty na fazie przewodzącej jony



Studium parametryczne: potencjał faz przewodzących jony dla różnych grubości elektrolitu

Nagrody Narodowego Centrum Nauki dla użytkowników zasobów obliczeniowych ACK Cyfronet AGH

Główna misja Cyfronetu to służba środowisku naukowemu, aktywne wspieranie m.in. indywidualnych badaczy poprzez udostępnianie im zasobów teleinformatycznych Centrum: od mocy obliczeniowej superkomputerów, poprzez pamięci masowe, aż po superszybką sieć komputerową i przydatne narzędzia informatyczne. Bardzo cieszy nas fakt, że wśród szerokiego zakresu tematów i prac badawczych realizowanych przy pomocy superkomputerów ACK Cyfronet AGH znalazły się również takie, które zostały docenione i nagrodzone przez Narodowe Centrum Nauki. Sukcesy naukowe użytkowników infrastruktury Cyfronetu najlepiej świadczą bowiem o jej użyteczności jako narzędzia prowadzenia prac naukowo-badawczych.



NARODOWE CENTRUM NAUKI

Nagroda Narodowego Centrum Nauki to wyróżnienie dla młodych uczonych za znaczące osiągnięcia naukowe, dokonane w ramach badań podstawowych prowadzonych w polskiej jednostce naukowej, udokumentowane publikacjami afiliowanymi w takiej jednostce. Nagrodę ustanowiła Rada Narodowego Centrum Nauki w lutym 2013 r. Wyróżnienie przyznawane jest w trzech obszarach badawczych: 1) nauki humanistyczne, społeczne i o sztuce, 2) nauki o życiu, 3) nauki ścisłe i techniczne. Wielokrotnie laureaci tych nagród podkreślają rolę zasobów obliczeniowych Cyfronetu w osiągnięciu wyników badań.

Nagroda Narodowego Centrum Nauki dla dr hab. Michała Tomzy z Wydziału Fizyki Uniwersytetu Warszawskiego*

W dniu 14 października 2020 laureatem Nagrody w obszarze nauk ścisłych i technicznych został dr hab. Michał Tomza. Nagroda została przyznana za opis teoretyczny oddziaływań i zderzeń pomiędzy ultrazimnymi atomami, jonami i cząsteczkami.

Dr hab. Michał Tomza jest fizykiem i chemikiem teoretycznym, adiunktem w Instytucie Fizyki Teoretycznej Wydziału Fizyki Uniwersytetu Warszawskiego. Specjalizuje się w kwantowym opisie materii w ultraniskich temperaturach, w tym teorii oddziaływań i zderzeń ultrazimnych atomów,

jonów i cząsteczek. W ramach staży naukowych prowadził badania naukowe na Uniwersytecie w Granadzie, Uniwersytecie Kolumbii Brytyjskiej w Vancouver, Uniwersytecie Harvarda, Uniwersytecie Kolorado w Boulder, Uniwersytecie Kalifornijskim w Santa Barbara oraz Uniwersytecie w Innsbrucku. W 2017 r. stworzył własną grupę badawczą na Uniwersytecie Warszawskim, na której działanie zdobył granty naukowe w konkursach SONATA i OPUS Narodowego Centrum Nauki oraz Homing i First Team z Fundacji na rzecz Nauki Polskiej.

Dr Tomza jest opiekunem naukowym kilkunastu magistrantów i doktorantów. Współpracuje blisko z grupami doświadczalnymi z Niemiec, Holandii i Szwajcarii. W ostatnim czasie jego zainteresowania naukowe skupiają się na ultrazimnych mieszaninach jonów z atomami.

– W mojej pracy naukowej badam właściwości, dynamikę oraz potencjalne zastosowania materii w ultraniskich temperaturach – mówi dr hab. Michał Tomza. – Ultrazimna materia to atomy, jony lub cząsteczki o temperaturze bliskiej zera bezwzględnego, czyli znacznie mniejszej niż np. temperatura przestrzeni kosmicznej. W takich warunkach bardzo wyraźnie ujawnia się kwantowa natura świata, dlatego moje badania pozwalają uzyskać nowy wgląd w teorię materii oraz oddziaływań pomiędzy materią i światłem – tematy ważne zarówno z poznawczego, jak i technologicznego punktu widzenia. W mikroskali, na poziomie atomowym, cały otaczający nas świat opisany jest przez prawa fizyki kwantowej, a szereg wykorzystywanych obecnie urządzeń, takich jak laser, tomograf czy GPS, nie mogłoby istnieć, gdyby nie efekty kwantowe. Obecnie stoimy także u progu narodzin nowych technologii kwantowych, które w bezpośredni sposób będą wykorzystywać podstawowe prawa natury do zagwarantowania maksymalnej czułości detektorów, maksymalnej wydajności komputerów, czy maksymalnego bezpieczeństwa szyfrowania. Moim celem jest rozwijanie koncepcji oraz metod teoretycznych do kwantowego opisu i odkrywania właściwości materii w ultraniskich temperaturach. Wyniki teoretyczne, które otrzymuję, są wykorzystywane w nowych eksperymentach na pograniczu fizyki i chemii oraz przyczyniają się do rozwoju nowych technologii kwantowych.

Dr hab. Michał Tomza wspiera swoje badania obliczeniami na udostępnianym przez Cyfronet superkomputerze Prometheus, w ramach grantu obliczeniowego plgquantmol4: „Symulacje ultrazimnych kwantowych wielociałowych układów molekularnych”.

Serdecznie gratulujemy Laureatowi i życzymy dalszych sukcesów!

**Materiały prasowe za <https://www.ncn.gov.pl/>*



CDC CYBER 72



Convex C3840



Exemplar SPP1600/XA

- 1973** Powstanie CYFRONETu
- 1975** Instalacja komputera CDC CYBER 72 w Centrum
- 1990** Instalacja pierwszego w Krakowie węzła sieci EARN / BITNET (komputer IBM 4381)
- 1991** Instalacja Convex 120 – pierwszego komputera wektorowego w Europie Środkowo-Wschodniej. Pierwsze internetowe połączenie międzymiastowe z Warszawą. Początek budowy Miejskiej Sieci Komputerowej w Krakowie
- 1994** Uruchomienie połączenia internetowego z Warszawą o prędkości 2 Mb/s
- 1996** Komputer Exemplar SPP1600/XA umieszczony na prestiżowej liście TOP500. Instalacja pierwszej automatycznej biblioteki taśmowej ATL 2640
- 1997** Uruchomienie podsieci komunikacyjnej ATM w Miejskiej Sieci Komputerowej. Przyłączenie Centrum do krajowej sieci POL-34
- 1998** Instalacja komputera SGI Origin2000
- 2000** Zwiększenie przepustowości połączenia sieciowego Centrum do 155 Mbps
- 2002** Instalacja komputera RackSaver PC w ramach projektu CrossGrid
- 2003** Instalacja pierwszego w Polsce komputera HP Integrity SuperDome
- 2005** Instalacja macierzy dyskowej HP Storage Works XP12000. Zwiększenie przepustowości połączenia sieciowego Centrum do 622 Mbps
- 2006** Instalacja macierzy dyskowej HP Storage Works EVA 8000 i superkomputera SGI ALTIX 3700 (Baribal) o teoretycznej mocy obliczeniowej 0,8 TFlops

- 2007** Podpisanie umowy Konsorcjum PL-Grid dotyczącej ogólnopolskiej infrastruktury gridowej PL-Grid.
Instalacja superkomputera SGI ALTIX 4700 z modulem akceleracyjnym SGI RASC.
Instalacja serwerów IBM BladeCenter HS21 (6,2 TFlops).
Instalacja macierzy dyskowej HP EVA 8100
- 2008** Rozbudowa konfiguracji superkomputera SGI ALTIX 3700 (do 1,5 TFlops).
Rozpoczęcie eksploatacji łącz sieci szkieletowej w technologii 10 Gbps Ethernet.
Połączenie MSK w kierunku Warszawy i Bielska-Białej łączami sieci PIONIER, każde o przepływności 2x10 Gbps.
Instalacja superkomputera Zeus (HP Cluster Platform 3000 BL), posiadającego 2 048 rdzeni
- 2009** Rozpoczęcie projektu „Infrastruktura Informatycznego Wspomagania Nauki w Europejskiej Przestrzeni Badawczej – PLGrid”
- 2010** Po rozbudowie konfiguracji do 9 544 rdzeni Intel Xeon superkomputer Zeus zostaje wpisany na listę TOP500 na 161 pozycji
- 2011** Wdrożenie wydajnego sieciowego systemu współdzielenia plików dla infrastruktury obliczeniowej, bazującego na rozwiązaniach firmy Hitachi Data Systems.
Całkowita pojemność zasobów dyskowych przekroczyła 2 PB.
Rozbudowa konfiguracji superkomputera Zeus do 12 032 rdzeni Intel Xeon.
Zeus zajmuje 81 miejsce na liście TOP500
- 2012** Rozpoczęcie realizacji projektu PLGrid Plus – dziedzinowo zorientowane usługi i zasoby w infrastrukturze PL-Grid.
W kwietniu, ScaleMP, wiodący dostawca wirtualnych rozwiązań dla superkomputerów, ogłosił „Zeusa-vSMP” największym systemem v-SMP w Europie.
Zeus w pierwszej setce rankingu TOP500.
Połączenie MSK w kierunku Rzeszowa łączem sieci PIONIER o przepływności 2x10 Gbps



SGI Origin2000



SGI ALTIX 3700



HP Cluster Platform 3000 BL



Medal jubileuszowy



Hala Maszyn



Superkomputer Prometheus

- 2013** Po rozbudowie do 25 468 rdzeni, superkomputer Zeus osiągnął teoretyczną moc obliczeniową 374 TFlops.
Wybicie medalu z okazji jubileuszu 40-lecia ACK Cyfronet AGH
- 2014** Oddano do użytku nową Halę Maszyn.
Rozpoczęcie realizacji dwóch nowych projektów: PLGrid NG i PLGridCore.
Połączenie MSK w kierunku Katowic łączem sieci PIONIER o przepływności 2x10 Gbps
- 2015** Instalacja superkomputera Prometheus (41 472 rdzenie), który na liście TOP500 zajmuje 49 miejsce (edycja lipcowa), a następnie (po rozbudowie do 53 568 rdzeni) 38 miejsce (edycja listopadowa).
Po raz pierwszy w historii polskiej informatyki dwa superkomputery z jednego ośrodka (Zeus i Prometheus) znalazły się jednocześnie na prestiżowej liście TOP500.
Oddano do użytku nowe, zapasowe centrum danych.
Rozpoczęto realizację projektów INDIGO-DataCloud, EGI-Engage, EPOS-IP i PRACE 4IP.
Łączność pomiędzy Centrami Komputerów Dużej Mocy w Polsce (Gdańsk, Kraków, Poznań, Warszawa i Wrocław) realizowana jest z przepustowością 2x100 Gbps
- 2016** Prometheus zajmuje 48 miejsce (edycja czerwcowa) i 59 miejsce (edycja listopadowa) na liście TOP500
- 2017** Prometheus zajmuje 71 miejsce (edycja czerwcowa) i 77 miejsce (edycja listopadowa) na liście TOP500.
Dalszy dynamiczny rozwój Centrum, w tym powstanie 6 nowych laboratoriów.
Rozpoczęto realizację projektów Sat4Envi, Gliomed, EPOS-PL oraz eXtreme DataCloud
- 2018** Prometheus (53 604 rdzenie, 2,4 PFlops) zajmuje 103 miejsce (edycja czerwcowa) i 131 miejsce (edycja listopadowa) na liście TOP500.
Rozpoczęcie realizacji projektów EOSC-Hub i PRIMAGE

2019 Cyfronet reprezentuje Polskę w konsorcjum LUMI, złożonym z ośmiu krajów, które będą wspólnie budować jeden z najszybszych europejskich superkomputerów.

Prometheus zajmuje 174 miejsce (edycja czerwcową) i 241 miejsce (edycja listopadową) na liście TOP500.

Prezentacja Cyfronetu na stoisku wystawowym w trakcie konferencji ISC'19.

Cyfronet udostępnia nowy system obliczeniowy dla badań z użyciem metod sztucznej inteligencji o mocy ponad 4 PFlops dla operacji tensorowych i 256 TFlops dla standardowych obliczeń.

Rozpoczęto realizację projektów PRACE-LAB, PRACE-6IP, SANO, EOSC-Synergy i EOSC Enhance



2020 Wśród strategicznych infrastruktur wpisanych w styczniu 2020 r. na Polską Mapę Infrastruktury Badawczej znajdują się dwa przedsięwzięcia zgłoszone przez ACK Cyfronet AGH jako inicjatora i koordynatora konsorcjum PLGrid: *Narodowa Infrastruktura Superkomputerowa dla EuroHPC* oraz *Narodowa Infrastruktura Chmurowa PLGrid dla EOSC*.

Prometheus (53 748 rdzeni, 2,65 PFlops) zajmuje 288 miejsce (edycja czerwcową) i 324 miejsce (edycja listopadową) na liście TOP500.

Superkomputer Prometheus wspomaga naukowców w walce z koronawirusem.

Rozpoczęto realizację projektów EPOS PL +, PRACE-LAB2, EPOS SP, PROTEUS-RS i EUROCC



2021 Uruchomiono superkomputer Ares o mocy obliczeniowej ponad 3,5 PFlops.

Prometheus i Ares zajmują odpowiednio 373 i 216 miejsce (edycja czerwcową) na liście TOP500.

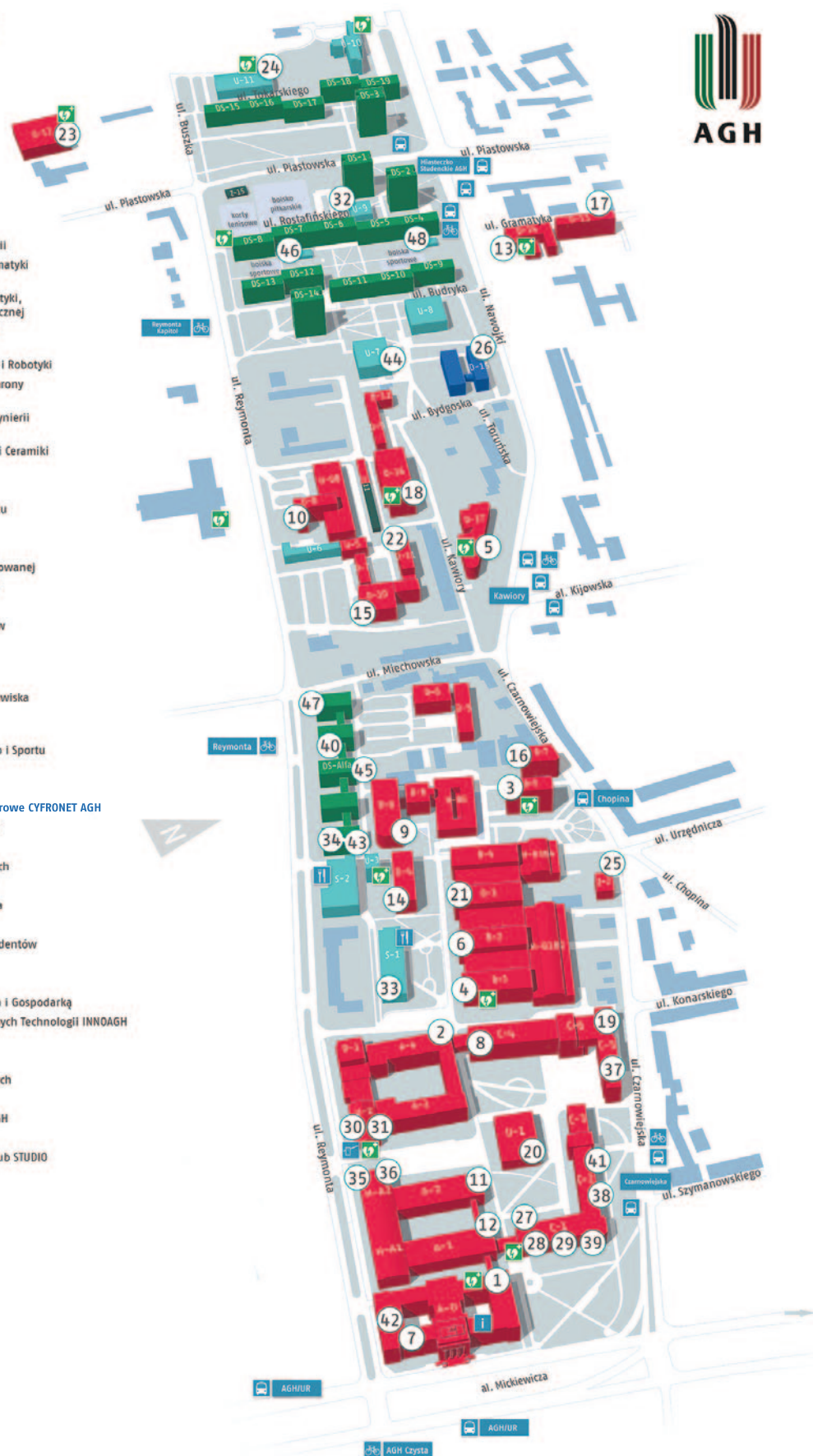
Uruchomiono nową wersję strony internetowej ACK Cyfronet AGH.

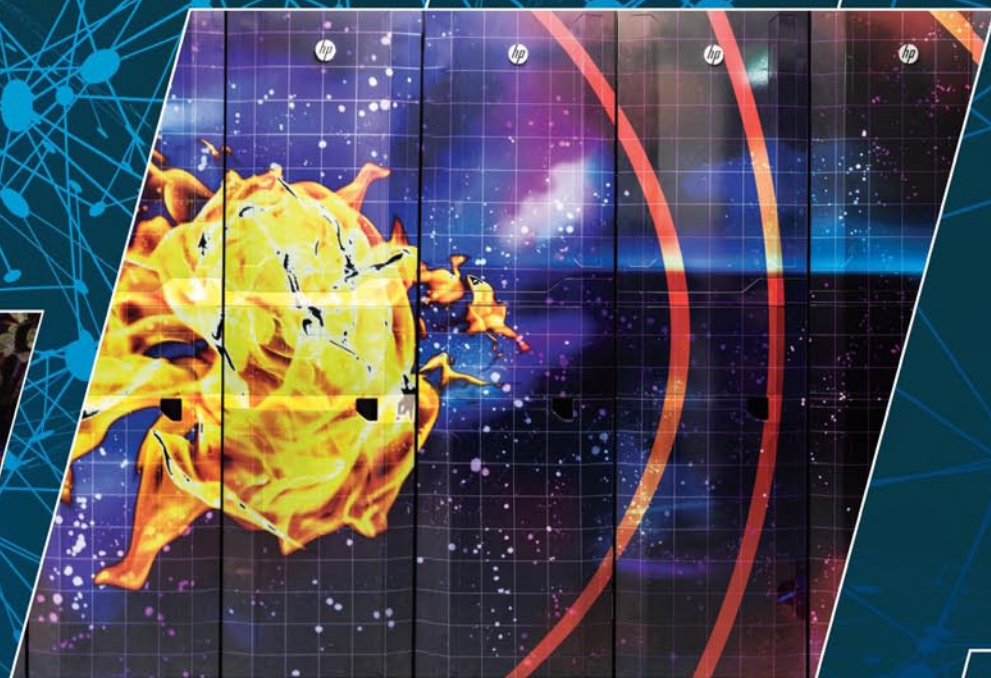
Rozpoczęto realizację projektów EuroHPC PL, PIONIER-LAB, KMD3, EGI ACE i EOSC Future.





1. Rektorat
2. Wydział Górnictwa i Geoinżynierii
3. Wydział Inżynierii Metali i Informatyki Przemysłowej
4. Wydział Elektrotechniki, Automatyki, Informatyki i Inżynierii Biomedycznej
5. Wydział Informatyki, Elektroniki i Telekomunikacji
6. Wydział Inżynierii Mechanicznej i Robotyki
7. Wydział Geologii, Geofizyki i Ochrony Środowiska
8. Wydział Geodezji Górniczej i Inżynierii Środowiska
9. Wydział Inżynierii Materiałowej i Ceramiki
10. Wydział Odlewnictwa
11. Wydział Metali Nieżelaznych
12. Wydział Wiertnictwa, Nafty i Gazu
13. Wydział Zarządzania
14. Wydział Energetyki i Paliw
15. Wydział Fizyki i Informatyki Stosowanej
16. Wydział Matematyki Stosowanej
17. Wydział Humanistyczny
18. Akademickie Centrum Materiałów i Nanotechnologii AGH
19. Centrum Energetyki AGH
20. Biblioteka Główna
21. Szkoła Ochrony i Inżynierii Środowiska im. Walerego Goetla
22. Studium Języków Obcych
23. Studium Wychowania Fizycznego i Sportu
24. Basen AGH
25. Centrum e-Learningu AGH
26. Akademickie Centrum Komputerowe CYFRONET AGH
27. Uczelniane Centrum Informatyki
28. Dział Nauczania
29. Centrum Studentów Zagranicznych
30. Centrum Dydaktyki AGH
31. Uczelniana Komisja Rekrutacyjna
32. Miasteczko Studenckie AGH
33. Uczelniana Rada Samorządu Studentów
34. Centrum Karier AGH
35. Centrum Transferu Technologii
36. Dział Współpracy z Administracją i Gospodarką
37. Krakowskie Centrum Innowacyjnych Technologii INNOAGH
38. Centrum Obsługi Projektów
39. Dział Współpracy z Zagranicą
40. Biuro ds. Osób Niepełnosprawnych
41. Muzeum AGH
42. Muzeum Geologiczne WGGiOŚ AGH
43. Wydawnictwa AGH
44. Akademickie Centrum Kultury Klub STUDIO
45. Klub Studencki Gwarek
46. Klub Studencki Zaścianek
47. Klub Studencki Karlik
48. Klub Studencki Filutek





ISBN 978-83-61433-38-5

Akademickie Centrum Komputerowe
CYFRONET AGH
ul. Nawojki 11

30-950 Kraków, P.O.Box 386
telefon: +48 12 632 33 55
fax: +48 12 633 80 54

cyfronet@cyfronet.pl
www.cyfronet.pl