



Akademickie Centrum Komputerowe
CYFRONET AGH



2025



ACK CYFRONET AGH jest jednostką wiodącą, ustanowioną przez Komitet Badań Naukowych, w zakresie eksploatacji i rozbudowy Komputerów Dużej Mocy (KDM) oraz Akademickiej Miejskiej Sieci Komputerowej (MSK). CYFRONET jest koordynatorem Programu PLGrid oraz uznanym przez NCBiR Centrum Kompetencji w zakresie rozproszonych infrastruktur obliczeniowych typu gridowego i chmurowego.

Szanowni Państwo!

Minął kolejny rok intensywnych prac mających na celu dostarczenie naszym Użytkownikom najnowszych technologii wspierających rozwój polskiej nauki, innowacyjnej gospodarki i społeczeństwa informacyjnego.

Wraz z uruchomieniem Heliosa, który przejął od Atheny tytuł najszybszego superkomputera w Polsce, znacznie rozszerzyliśmy możliwości wykonywania zaawansowanych zadań obliczeniowych, w tym na potrzeby kształtowania sztucznej inteligencji. Helios posiada 35 PetaFlopsów teoretycznej mocy obliczeniowej, pochodzącej głównie z procesorów graficznych, co pozwoliło między innymi na efektywne wsparcie tworzenia polskich modeli językowych Bielik i PLLuM. Na listopadowej liście Top500 najszybszych superkomputerów świata Helios został sklasyfikowany dwukrotnie, w tym jako partycja GPU na wysokim, 69 miejscu. Na tej samej liście odnotowano również Athenę i Aresa, co sprawiło, że połowa z ośmiu pozycji zajętych przez polskie maszyny to superkomputery działające w Cyfronecie.

Warto jednak zaznaczyć, że dbamy nie tylko o rozwój możliwości obliczeniowych, ale także o jak najefektywniejsze ich wykorzystanie, biorąc pod uwagę również kwestie środowiskowe. Potwierdzeniem tego faktu jest przede wszystkim zajęcie przez partycję GPU Heliosa siódmego miejsca na prestiżowej liście Green500 najbardziej efektywnych energetycznie superkomputerów świata. Mając jeszcze na uwadze, że na poprzedniej edycji tej listy Helios znalazł się na 3. miejscu, a w 2022 roku Athenę odnotowano na 10 pozycji, możemy śmiało powiedzieć, że budowane przez nas superkomputery należą do światowej czołówki najbardziej efektywnych maszyn, a nasze kompetencje w zakresie optymalizacji systemów i oprogramowania są na światowym poziomie.

Jednocześnie, w kolejny rok działalności wkraczamy aktywnie uczestnicząc w wielu innych inicjatywach. Cyfronet organizuje kolejne konkursy na roczny dostęp do obliczeń na superkomputerze LUMI, będącym pod koniec 2024 roku trzecim superkomputerem w Europie pod względem mocy obliczeniowej. Dodatkowo, dzięki reprezentowaniu Polski w międzynarodowym konsorcjum LUMI-Q, superkomputer Helios będzie bezpośrednio połączony z budowanym w Czechach superkomputerem kwantowym.

Wszystkie wspomniane zasoby są dostępne w ramach zainicjowanej i koordynowanej przez Cyfronet infrastruktury PLGrid. Podjęta przed laty strategia unifikacji w ramach infrastruktury PLGrid dostępu do różnych zasobów przyniosła wiele korzyści dla użytkowników. Po zalogowaniu się do Portalu PLGrid można w łatwy sposób wybierać spośród wielu maszyn, środowisk wirtualnych i pakietów oprogramowania. Niezwykle cennego wsparcia udziela zespół Helpdesku czyli Centrum Operacyjne, a ponadto Cyfronet organizuje dla użytkowników szkolenia z efektywnego wykorzystania udostępnianych narzędzi. W tym zakresie warto wspomnieć o koordynowaniu przez Cyfronet Narodowego Centrum Kompetencji Obliczeniowych, tworzonego wspólnie z pozostałymi centrami KDM w Polsce, które pełni rolę punktu kontaktowego w zakresie obliczeń dla naukowców, startupów, małych i średnich przedsiębiorstw, dużych firm oraz administracji publicznej.

W 2024 roku rozpoczęliśmy kolejny z serii projektów modernizujących PLGrid: Narodowa Infrastruktura Chmurowa PLGrid dla EOSC – PLGrid ICON, który ma na celu zbudowanie nowoczesnej infrastruktury chmurowej, umożliwiającej realizację prac badawczo-rozwojowych zgodnie z paradygmatem Otwartej Nauki. O jej rozwój dbamy również poprzez aktywne wspieranie rozbudowy środowiska EOSC, a także doskonalenie tzw. Cyfrowych Bliźniaków (Digital Twins), w tym w medycynie i naukach o Ziemi.

O tych, oraz o wielu innych działaniach Cyfronetu przeczytają Państwo na kolejnych kartach niniejszej publikacji, do czego z całego serca zapraszam. Jednocześnie bardzo dziękuję wszystkim Przyjaciółom i Użytkownikom Cyfronetu za współpracę i cenne wskazówki dotyczące dalszego rozwoju Centrum.

Z wyrazami szacunku

Marek Magryś

p. o. Dyrektor ACK Cyfronet AGH



p. o. Dyrektor

inż. Marek Magryś

Tel. +48 12 633-34-26

e-mail: Marek.Magrys@cyfronet.pl



Zastępca Dyrektora

**ds. Bezpieczeństwa i Innowacyjnej
Gospodarki**

mgr inż. Henryk Baniowski

Tel. +48 12 633 34 26

e-mail: Henryk.Baniowski@cyfronet.pl



Zastępca Dyrektora

ds. Infrastruktury Informatycznej

mgr inż. Karol Krawentek

Tel. +48 12 633 34 26

e-mail: Karol.Krawentek@cyfronet.pl



Pełnomocnik Dyrektora

ds. Ekonomicznych

mgr Emilia Piecuch

Tel. +48 12 633 34 26

e-mail: Emilia.Piecuch@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 34 26

e-mail: Angelika.Zaleska@cyfronet.pl



Rada Cyfronetu

Przewodniczący Rady

prof. dr hab. inż. Marek Gorgoń

Prorektor ds. Nauki AGH

mgr inż. Henryk Baniowski

Zastępca Dyrektora

ds. Bezpieczeństwa i Innowacyjnej Gospodarki

mgr inż. Karol Krawentek

Zastępca Dyrektora

ds. Infrastruktury Informatycznej

inż. Marek Magryś

p. o. Dyrektor ACK Cyfronet AGH

mgr Emilia Piecuch

Pełnomocnik Dyrektora

ds. Ekonomicznych

mgr Agnieszka Szymańska

Zastępca Dyrektora

ds. Administracyjnych

mgr Angelika Zaleska-Walterbach

Zastępca Dyrektora

ds. Ekonomiczno-Finansowych

prof. dr hab. inż. Kazimierz Wiatr

Kadencja 2025-2028

prof. dr hab. Armen Edigarian
Przewodniczący

Uniwersytet Jagielloński

mgr inż. Jakub Badowski

Instytut Nafty i Gazu
Państwowy Instytut Badawczy

p.o. Dyrektora Jacek Barski

Urząd Marszałkowski Województwa Małopolskiego

dr hab. Piotr Batys

Instytut Katalizy i Fizykochemii Powierzchni PAN

mgr Michał Furman

Uniwersytet Jagielloński

prof. dr hab. inż. Marek Gorgoń

Akademia Górniczo-Hutnicza

mgr inż. Jarosław Górski

Akademia Kultury Fizycznej

prof. dr hab. inż. Marek Kisiel-Dorohinicki

Akademia Górniczo-Hutnicza

ks. dr hab. Józef Kloch

Uniwersytet Papieski Jana Pawła II

mgr inż. Mariusz Kwinta-Pudółko

Politechnika Krakowska

prof. dr hab. Tadeusz Lesiak

Instytut Fizyki Jądrowej PAN

mgr inż. Wojciech Majka

Uniwersytet Jagielloński Collegium Medicum

mgr inż. Tomasz Masier

Instytut Ekspertyz Sądowych

dr hab. inż. Piotr Pokora

Uniwersytet Komisji Edukacji Narodowej

mgr inż. Mateusz Ponikiewicz

Uniwersytet Rolniczy

prof. dr hab. Dariusz Put

Uniwersytet Ekonomiczny

Dyrektor Paweł Schmidt

Urząd Miasta Krakowa

dr hab. Bartłomiej Struzik

Akademia Sztuk Pięknych

mgr Maciej Tomkiewicz

Sieć Badawcza Łukasiewicz
Krakowski Instytut Technologiczny

Akademickie Centrum Komputerowe CYFRONET AGH jest najdłużej działającym i jednym z największych centrów superkomputerowych i sieciowych w Polsce, z historią udostępniania zasobów superkomputerowych sięgającą 1975 roku.

ACK Cyfronet AGH jest od lat operatorem najszybszych superkomputerów w Polsce, wielokrotnie notowanych na światowej liście TOP500, a także systemów składowania danych o bardzo dużej pojemności. Dysponuje trzema centrami danych, własną siecią światłowodową oraz zapleczem technicznym, personelem i kompetencjami pozwalającymi na działanie 24 godziny na dobę, 365 dni w roku. Centrum pełni rolę administratora sieci MAN w Krakowie i jest ważnym węzłem sieci akademickiej PIONIER, która jest połączona z europejską siecią GÉANT.

Cyfronet jest organizatorem i liderem Konsorcjum PLGrid, konsolidującego krajowe zasoby obliczeniowe i zapewniającego szereg unikalnych usług obliczeniowych oraz usług wsparcia IT dla nauki, a także liderem Narodowego Centrum Kompetencji HPC, które pełni rolę punktu kontaktowego i dostępowego do HPC zarówno dla środowiska akademickiego, jak i innowacyjnych podmiotów gospodarki i administracji publicznej.

ACK Cyfronet AGH pełni rolę koordynatora w przedsięwzięciach wpisanych na Polską Mapę Infrastruktury Badawczej (PMIB): Narodowa Infrastruktura Superkomputerowa dla EuroHPC oraz Narodowa Infrastruktura Chmurowa PLGrid dla EOSC. Jest także członkiem konsorcjów realizujących programy wpisane na PMIB: CTA, EPOS, KMD, PIONIER-LAB, PRACE i Virgo, a dodatkowo dostarcza infrastrukturę obliczeniową Komputerów Dużej Mocy wraz z zasobami systemów pamięci masowej oraz infrastrukturę informatyczną Miejskiej Akademickiej Sieci Komputerowej w Krakowie (MAN) na potrzeby prowadzenia prac badawczych w innych projektach PMIB.

Cyfronet jest aktywnie zaangażowany w czołowe europejskie projekty związane z rozwojem technologii superkomputerowych i opartych na nich usług, w tym m.in: WLCG (The Worldwide LHC Computing Grid), EGI (Advanced Computing Services for Research), PRACE (Partnership for Advanced Computing in Europe), EuroHPC JU (The European High Performance Computing Joint Undertaking), EOSC (European Open Science Cloud), EPOS (European Plate Observing System), LUMI (The Large Unified Modern Infrastructure), LUMI AI, LUMI-Q, Digital Twins w naukach o Ziemi i zdrowiu.

Jednocześnie Cyfronet uczestniczy w wielu innych krajowych i międzynarodowych projektach badawczych i rozwojowych, w których wykorzystywane są zarówno zasoby sprzętowe, jak i unikalne doświadczenie w budowaniu i rozwijaniu zintegrowanych platform usługowych dla użytkowników naukowych. Działające w Cyfronecie Laboratoria skupiają specjalistów opracowujących narzędzia informatyczne wspierające rozwój nauki i technologii, między innymi w zakresie nowoczesnej diagnostyki medycznej i terapii.

Centrum współpracuje również z MŚP i dużymi firmami, aby umożliwić efektywne wdrażanie HPC (High Performance Computing) i AI (Artificial Intelligence) w logistyce, medycynie, przetwarzaniu obrazów satelitarnych, CFD, badaniach nad lekami i przemyśle motoryzacyjnym.

Ważnym aspektem działalności Cyfronetu jest organizacja wydarzeń o charakterze naukowym oraz specjalistycznych szkoleń i warsztatów związanych z HPC.

Certyfikat ISO 27001:2023

Na mocy audytu przeprowadzonego przez Polskie Centrum Badań i Certyfikacji S. A., w dn. 30 grudnia 2024 roku Akademickie Centrum Komputerowe CYFRONET AGH otrzymało certyfikat zgodności z normą PN-EN ISO/IEC 27001:2023-08 (ISO 27001).

Certyfikat potwierdza spełnienie wymagań normy w następujących zakresach:

1. usługi:
 - przetwarzania i składowania danych,
 - HPC,
 - chmury obliczeniowej,
 - sztucznej inteligencji,
 - kolokacji i centrum danych,
 - cyberbezpieczeństwa,
 - sieciowe,
 - transmisji danych,
2. wytwarzanie oprogramowania,
3. audyty bezpieczeństwa,
4. szkolenia i konsulting IT.



ACK Cyfronet AGH posiada Certyfikat zgodności z normą ISO 27001:2023

Przyznanie certyfikatu stanowi potwierdzenie, że ACK Cyfronet AGH spełnia wysokie wymagania międzynarodowych standardów.





Helios – nowe światło dla innowacji w polskiej nauce i gospodarce

Superkomputer Helios to system zainstalowany w ACK Cyfronet AGH, a powstały w wyniku prac realizowanych w koordynowanym przez Cyfronet projekcie Narodowa Infrastruktura Superkomputerowa dla EuroHPC – EuroHPC PL. Superkomputer został zbudowany według projektu Cyfronetu przez Hewlett-Packard Enterprise w oparciu o platformę HPE Cray EX4000 i składa się z trzech partycji obliczeniowych:

- partycja CPU wyposażona w 75.264 rdzenie obliczeniowe AMD Zen4 oraz 200 TB pamięci operacyjnej DDR5,
- partycja GPU wyposażona w 440 superczipów NVIDIA Grace Hopper GH200,
- partycja INT dla pracy interaktywnej, wyposażona w 24 akceleratory NVIDIA H100 i szybką lokalną pamięć NVMe.

Helios osiąga 35 PFlops teoretycznej mocy obliczeniowej.

Helios w liczbach	
Liczba rdzeni obliczeniowych	75 264
Liczba kart z procesorami graficznymi GPGPU	464
Moc obliczeniowa	35 PFlops
Ranking TOP500 najszybszych komputerów świata (listopad 2024)	69 lokata (partycja GPU)
Ranking Green500 najefektywniejszych energetycznie komputerów świata (listopad 2024)	7 lokata (partycja GPU)

Moc obliczeniowa Heliosa dla obliczeń AI to 1,8 ExaFlops!

Podsystem dyskowy Heliosa składa się z dwóch rodzajów systemów plików Lustre: *scratch* o pojemności 1,5 PB i prędkości ponad 1,8 TB/s oraz *project* o pojemności 16 PB i prędkości niemal 200 GB/s. Wszystkie komponenty superkomputera połączone są ze sobą siecią Slingshot o prędkości 200 Gb/s. Dzięki bezpośredniemu chłodzeniu cieczą partycji CPU i GPU, możliwe jest osiągnięcie bardzo niskiego wskaźnika PUE (Power Usage

Effectiveness) dla systemu, a w efekcie zwiększenie jego efektywności energetycznej i obniżenie kosztów eksploatacji. Dodatkowo, dzięki odzyskowi ciepła odpadowego produkowanego przez Heliosa, możliwe będzie wykorzystanie go do ogrzewania.



Athena – mocne wsparcie dla obliczeń naukowych

Athena osiąga teoretyczną moc obliczeniową ponad **7,7 PFlops**. Zainstalowany w 2021 roku w Cyfronecie system dostarcza polskiemu środowisku naukowemu zasoby obliczeniowe oparte na procesorach i akceleratorach GPGPU najnowszej generacji wraz z niezbędnym podsystemem składowania danych opartym na bardzo szybkich pamięciach flash. Konfiguracja Atheny obejmuje: 48 serwerów z procesorami AMD EPYC i 1 TB pamięci RAM (w sumie 6144 rdzenie obliczeniowe CPU) oraz 384 karty GPGPU NVIDIA A100.

Niezbędnym elementem umożliwiającym wykorzystanie tak dużej mocy obliczeniowej w efek-

tywny sposób jest zapewnienie wysokowydajnej sieci wewnętrznej superkomputera (Infiniband HDR o przepustowości 4x200 Gb/s na serwer) oraz bardzo szybkiego podsystemu dyskowego. Jest on zbudowany w oparciu o otwarte oprogramowanie Lustre, używane również w systemach dyskowych superkomputerów Ares i Prometheus, oraz dedykowane serwery dyskowe wyposażone w pamięci flash w standardzie NVMe. System został zainstalowany w istniejącym centrum danych Cyfronetu oraz zintegrowany z infrastrukturą PLGrid.

Architektura Atheny wychodzi naprzeciw potrzebom

użytkowników superkomputerów Cyfronetu, którzy wykorzystują infrastrukturę obliczeniową zarówno do wykonywania standardowych wysokowydajnych symulacji naukowych (HPC), jak i do aplikowania metod sztucznej inteligencji (AI) i uczenia maszynowego (ML) w badaniach z zakresu medycyny, farmakologii, biologii, chemii, fizyki oraz wielu innych dziedzin nauki. **Moc obliczeniowa Atheny do obliczeń AI to ponad 240 PFlops!**

Spodziewanym efektem dostarczenia specjalizowanych zasobów obliczeniowych Atheny będzie poszerzenie zakresu prowadzonych prac badawczych, możliwość podejmowania zaawansowanych symulacji i analiz oraz zwiększenie możliwości przetwarzania wpływających w sposób ciągły danych z laboratoriów z całego świata. Bezpośrednim spodziewanym efektem prowadzonych prac będą artykuły i opracowania naukowe, patenty, a w dalszej perspektywie nowatorskie rozwiązania, które mogą być podstawą dla powstawania nowych rozwiązań w gospodarce.

Athena w liczbach	
Liczba rdzeni obliczeniowych	6 144
Liczba kart z procesorami graficznymi GPGPU	384
Moc obliczeniowa	7,7 PFlops
Ranking TOP500 najszybszych komputerów świata (listopad 2024)	212 lokata



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, każdy wyposażony w 192 GB pamięci RAM,
- 256 serwerów, każdy posiadający 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 (listopad 2024)	490 lokata

Sumaryczna, teoretyczna wydajność części CPU to ponad 3,5 PFlops, a części GPU to ponad 500 TFlops. Aresa wspomaga 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 ważną składową zasobów obliczeniowych Cyfronetu poprzez dostarczenie nowszej gene-

racji procesorów i serwerów z większą ilością pamięci. Umożliwia 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 Data Center Podole, czyli w innej, geograficznie oddalonej lokalizacji niż Helios, gwarantuje zachowanie ciągłości świadczenia usług obliczeniowych w sytuacjach kryzysowych.





Prometheus – PetaFlops’owa moc



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 2,7 PFlops.

Prometheus w liczbach	
Liczba rdzeni obliczeniowych	53 748
Pamięć operacyjna	283,5 TB
Liczba kart z procesorami graficznymi GPGPU	176
Moc obliczeniowa	2,7 PFlops

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 temperaturach, 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.

Architektura superkomputera Prometheus



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.

Od momentu instalacji w roku 2015, Prometheus nieprzerwanie notowany był na liście TOP500 – łącznie aż 15 razy, zajmując wysokie miejsca, a najwyżej **38.** pozycję. Prometheus 11-krotnie był najszybszym superkomputerem w Polsce.

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.

Klaster technologii przyszłości Faeton

Faeton osiąga moc obliczeniową 288 TFlops i zbudowany jest z 64 serwerów obliczeniowych, z których każdy jest wyposażony w dwa procesory Intel Xeon Platinum 8352s wspierające szyfrowanie pamięci aplikacji, 1 TB pamięci operacyjnej oraz adaptery sieci 100 Gb/s Ethernet o niskich opóźnieniach. Dodatkowo, 4 serwery obliczeniowe zostały wyposażone w 8 TB pamięci SCM Intel Optane. W skład Faetona wchodzi także serwery usługowe oraz 12 serwerów dyskowych, oferujących ponad 1 PB pamięci dyskowej NVMe oraz 12 TB pamięci SCM. Taka konfiguracja stanowi doskonałe środowisko do rozwoju innowacyjnego oprogramowania, szczególnie w zakresie analityki danych (data science) oraz uruchamiania aplikacji w środowisku chmurowym o dużej wydajności i wysokim poziomie bezpieczeństwa.

Faeton będzie używany w celu weryfikacji możliwości stosowania nowych technologii, w szczególności pamięci typu SCM (Storage Class Memory) w aplikacjach obliczeniowych. Instalacja takiego systemu umożliwi prowadzenie prac nad zastosowaniem technologii przyszłości w aplikacjach polskich naukowców i przedsiębiorstw, dając im przewagę konkurencyjną poprzez udostępnienie platformy wykorzystującej rozwiązania na wczesnym etapie rozwoju, z dużym potencjałem innowacyjności.



Cztery superkomputery Cyfronetu na liście TOP500 superkomputerów o najwyższej mocy obliczeniowej



W 2024 roku, po raz pierwszy w historii na liście TOP500 (edycja czerwcową) znalazły się jednocześnie 4 superkomputery z jednego polskiego centrum obliczeniowego. Były to działające w Cyfronecie: Helios GPU (55 lokata), Athena (177), Helios CPU (305) i Ares (442). Sukces ten został powtórzony w listopadzie 2024, kiedy te same 4 komputery ponownie pojawiły się na liście TOP500, zajmując odpowiednio 69, 212, 348 i 490 miejsce.

Historia obecności zainstalowanych w Cyfronecie maszyn na liście TOP500 sięga roku 1996, kiedy to komputer SPP1200/XA-32 zajął 408 miejsce. Po dłuższym okresie bez notowań, w 2010 roku na listę trafił Zeus i pozostał tam aż do 2015 roku. Od tamtej pory do 2022 roku Cyfronet był w zestawieniu reprezentowany przez Prometheusa, do którego w 2021 dołączył Ares, w 2022 roku Athena, a w roku 2023 Helios.

Superkomputer Zeus

- 2010 - VI, 161 miejsce, 55 TFlops
- 2010 - XI, 85 miejsce, 105 TFlops
- 2011 - VI, **80** miejsce, 124 TFlops
- 2011 - XI, 88 miejsce, 162 TFlops
- 2012 - VI, 89 miejsce, 271 TFlops
- 2012 - XI, 106 miejsce, 357 TFlops
- 2013 - VI, 114 miejsce, 374 TFlops
- 2013 - XI, 146 miejsce, 374 TFlops
- 2014 - VI, 176 miejsce, 374 TFlops
- 2014 - XI, 211 miejsce, 374 TFlops

Superkomputery Prometheus i Zeus

- 2015 - VII, 49 i 269 miejsce, 1659+374 TFlops
- 2015 - XI, **38** i 387 miejsce, 2399+374 TFlops

Superkomputer Prometheus

- 2016 - VI, 49 miejsce, 2399 TFlops
- 2016 - XI, 60 miejsce, 2399 TFlops
- 2017 - VI, 72 miejsce, 2399 TFlops
- 2017 - XI, 78 miejsce, 2399 TFlops
- 2018 - VI, 103 miejsce, 2399 TFlops
- 2018 - XI, 131 miejsce, 2399 TFlops
- 2019 - VI, 174 miejsce, 2399 TFlops
- 2019 - XI, 241 miejsce, 2399 TFlops
- 2020 - VI, 288 miejsce, 2399 TFlops
- 2020 - XI, 324 miejsce, 2399 TFlops

Superkomputery Ares i Prometheus

- 2021 - VI, **216** i 373 miejsce, 3510+2399 TFlops
- 2021 - XI, 267 i 440 miejsce, 3510+2399 TFlops

Superkomputery Athena, Ares i Prometheus

- 2022 - VI, **105**, 290 i 475 miejsce, 7709+3510+2399 TFlops

Superkomputery Athena i Ares

- 2022 - XI, 113 i 323 miejsce, 7709+3510 TFlops
- 2023 - VI, 123 i 362 miejsce, 7709+3510 TFlops

Superkomputery Athena, Helios CPU i Ares

- 2023 - XI, 154, **290** i 403 miejsce, 7709+3400+3510 TFlops

Superkomputery Helios GPU, Athena, Helios CPU i Ares

- 2024 - VI, **55**, 177, 305 i 442 miejsce, 30,44+7,71+3,35+3,51 PFlops
- 2024 - XI, 69, 212, 348 i 490 miejsce, 30,44+7,71+3,35+3,51 PFlops

Superkomputery z Cyfronetu na liście TOP500

Czołówka światowej wydajności energetycznej

W 2024 roku cztery superkomputery Cyfronetu znalazły się również na liście Green500 – superkomputerów najbardziej ekologicznych, czyli o najwyższej wydajności energetycznej, obliczanej jako współczynnik ilości operacji zmiennoprzecinkowych na sekundę (mocy obliczeniowej superkomputera) do poboru energii: Gflops/W. Ogromnym sukcesem było zajęcie przez partycję GPU Heliosa **3 miejsca** w czerwcowej edycji listy oraz 7 miejsca w edycji listopadowej. Oznacza to, że Helios został najbardziej efektywnym energetycznie superkomputerem z pierwszej setki listy TOP500. Ta pozycja świadczy o doskonałym stosunku udostępnianej mocy obliczeniowej do poboru energii elektrycznej. Tak wysoka efektywność energetyczna oznacza, że Helios oferuje więcej mocy obliczeniowej za każdą skonsumowaną kilowatogodzinę niż mniej efektywne systemy, a więc obliczenia z jego wykorzystaniem są nie tylko tańsze niż w przypadku innych maszyn, ale także w mniejszym stopniu wpływają na środowisko. W 2024 roku nie zabrakło na liście Green500 pozostałych superkomputerów z Cyfronetu: Atheny, Heliosa CPU i Aresa, które zajęły odpowiednio 32, 95 i 129 miejsce w edycji czerwcowej oraz 53, 113 i 145 miejsce w edycji listopadowej.



Wykorzystanie superkomputerów

Krakowskie superkomputery są częścią europejskiej chmurowej i gridowej infrastruktury w ramach Europejskiej Infrastruktury Gridowej (EGI). Jednocześnie, 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 superkomputerów przez naukowców. Dedykowane środowiska obliczeniowe 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:

- badanie właściwości spektralnych związków chemicznych,
- rozwój modeli sztucznej inteligencji do zadań segmentacji zmian na obrazach CT płuc oraz MRI wątroby,
- detekcja pieszych na podstawie sygnału z kamery zdarzeniowej, z użyciem kwantyzowanych sieci neuronowych,
- predykcja fal grawitacyjnych za pomocą uczenia maszynowego,
- modelowanie separacji faz białek i polipeptydów,
- analiza danych meteorologicznych z wykorzystaniem uczenia maszynowego,
- modelowanie właściwości polimerów i nanomateriałów,
- symulacje dynamiki molekularnej elektrolitów,
- zastosowanie sztucznej inteligencji we wsparciu procesu diagnostycznego w medycynie weterynaryjnej.

Szeroki zakres tematów badawczych jest dowodem ciągle rosnącej liczby naukowców, którzy są świadomi korzyści płynących z wykorzystania superkomputerów. Z ich pomocą można uzyskać wyniki złożonych symulacji wiele, wiele razy szybciej niż na zwykłym komputerze stacjonarnym. Obliczenia, które przy użyciu pojedynczych komputerów często zajęłyby wiele 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 krakowskim superkomputerom polscy uczeni mogą uczestniczyć w ważnych międzynarodowych projektach nauko-

Rok	Liczba zadań	Czas CPU w latach
Superkomputery Cyfronetu		
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	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
2021	5 549 582	43 409
2022	6 227 244	48 716
2023	11 468 532	52 722

wych, w tym w eksperymentach z listy ESFRI i PMIB: CTA, LOFAR, EPOS, Wielkiego Zderzacza Hadronów w CERNie oraz fal grawitacyjnych w detektorach LIGO i VIRGO.

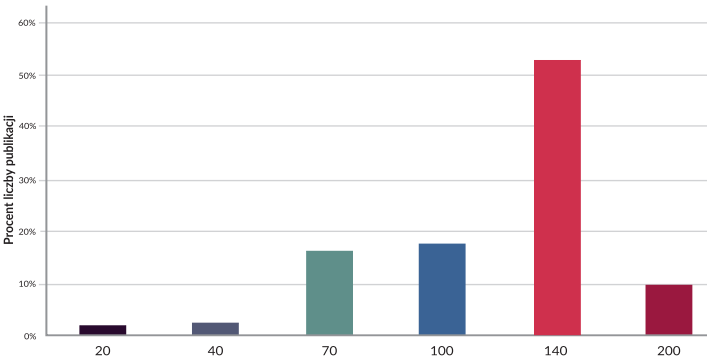
Oczywiście, nawet najwyższe lokaty w zestawieniu TOP500, czy też najnowocześniejsze technologie wykorzystane do budowy komputerów dużej mocy nie oddają w pełni znaczenia tego rodzaju zasobów obliczeniowych dla polskiego środowiska naukowego. O użyteczności superkomputerów udostępnianych przez ACK Cyfronet 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.

Poziom naukowy zadań zrealizowanych z wykorzystaniem infrastruktury udostępnianej przez ACK Cyfronet AGH jest bardzo wysoki. Dowodem tego są wyniki prac naukowo-badawczych, prowadzonych w roku 2023 przy wykorzystaniu tej infrastruktury, które zostały zaprezentowane w wielu publikacjach.

Punktacja artykułów naukowych według Ministerstwa, powstałych w 2023 roku dzięki zasobom ACK Cyfronet AGH



Kompleksowa infrastruktura wydajnego i bezpiecznego składowania danych cyfrowych

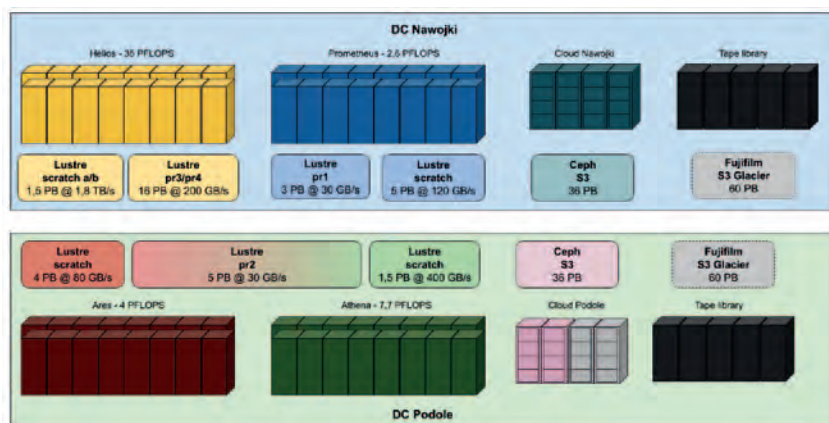
Obserwowane obecnie zjawisko lawinowo rosnącej ilości wytwarzanych informacji w postaci cyfrowej dotyczy także środowiska naukowego. Dostęp do bardzo wydajnych superkomputerów umożliwia podejmowanie analiz wielkoskalowych problemów badawczych, w efekcie czego powstają olbrzymie zbiory danych. Wymagają 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 widoczny jest wzrost oczekiwań w zakresie dostępnej pojemności, szybkości oraz dodatkowych funkcjonalności zasobów pamięci masowej, spowodowany 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 – 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 niższej 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 lub 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 i taśmowych Cyfronetu przekracza **270 PB**.

Pamięć masowa dla superkomputerów

Odpowiednie skorelowanie projektu architektury infrastruktury obliczeniowej z właściwym doбором 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 miliardy plików o rozmiarach sięgających 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. Zasoby systemu składowania danych Cyfronetu znajdują się w dwóch lokalizacjach.



W superkomputerach stosuje się m.in. wydajną przestrzeń tymczasową, tzw. *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 zagregowana 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 Cyfronetie wszystkie superkomputery korzystają z przestrzeni *scratch* realizowanych przez Lustre. W przypadku Prometheusa przestrzeń ta charakteryzuje się pojemnością 5 PB i szybkością 120 GB/s. Ares dysponuje przestrzenią o pojemności 4 PB i szybkością 80 GB/s. W obu tych komputerach przestrzeń *scratch* jest realizowana przy pomocy dysków mechanicznych. W przypadku Atheny i Heliosa dane użytkowników przechowywane są na dyskach półprzewodnikowych. Zastosowanie tego typu rozwiązania znacząco zwiększa wydajność systemu. Pojemność tego typu przestrzeni dla Atheny wynosi 1,5 PB oraz osiąga przepustowość 400 GB/s. W przypadku Heliosa przestrzeń *scratch* charakteryzuje się pojemnością 1,5 PB i szybkością 1,8 TB/s.

Większa część zasobów pamięci dyskowej Cyfronetu jest przeznaczona dla potrzeb użytkowników usług dziedzinyowych rozwijanych w programie PLGrid. Infrastruktura PLGrid oferuje dedykowaną przestrzeń roboczą dla grup korzystających z usług dziedzinyowych – funkcjonalność niezbędną do umożliwienia współpracy między naukowcami pracującymi w rozproszonych geograficznie lokalizacjach. Funkcjonalność ta jest realizowana za pomocą systemu plików Lustre. Maksymalna pojemność zasobu */pr1* w superkomputerze Prometheus to 5 PB, a sumaryczna szybkość operacji zapisu i odczytu sięga 30 GB/s. W przypadku Heliosa zasoby */pr3* i */pr4* charakteryzują się łączną pojemnością 16 PB i szybkością 200 GB/s.

Dodatkowymi zasobami do przechowywania zasobów danych użytkowników oraz projektów w Cyfronetie jest system obiektowego przechowywania danych. Jest on oparty o oprogramowanie CEPH. Dane w tym systemie dostępne są poprzez protokół S3 oparty na REST API i znajdują się w unikalnych globalnie kontenerach (bucket), w których użytkownicy przechowują swoje dane w formie obiektów.

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 Cyfronetie przekracza 2 PB.

Aktualnie całkowita udostępniana pojemność dyskowa eksploatowana przez ACK Cyfronet AGH wynosi około **150 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 ponad 9 tysięcy slotów na taśmy magnetyczne standardu LTO oraz 44 napędy generacji 6, 7 i 9. Pojedynczy nośnik magnetyczny LTO-9 posiada pojemność fizyczną 18 TB i pozwala na zapis z prędkością sięgającą 400 MB/s. 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.

Obecnie, całkowita pojemność zasobów taśmowych w Cyfronecie przekracza **120 PB**.



PLGrid – zaawansowane rozwiązania obliczeniowe dla polskiej nauki i gospodarki

PLGrid to koordynowana przez ACK Cyfronet AGH nowoczesna infrastruktura obliczeniowa, zbudowana i rozwijana dla wsparcia polskich sektorów nauki, gospodarki i administracji publicznej. Dzięki zaawansowanym zasobom superkomputerowym, kwantowym i chmurowym umożliwia prowadzenie prac badawczo-rozwojowych na najwyższym, światowym poziomie, wspierając tym samym innowacje i rozwój nowych technologii w wielu dziedzinach nauki i gospodarki.

Konsorcjum PLGrid

Początki infrastruktury PLGrid sięgają stycznia 2007 roku, kiedy to z inicjatywy ACK Cyfronet AGH powołano Konsorcjum PLGrid. W jego skład wchodzi największe polskie centra obliczeniowe:

- Akademickie Centrum Komputerowe Cyfronet AGH w Krakowie (lider konsorcjum),
- Centrum Informatyczne Trójmiejskiej Akademickiej Sieci Komputerowej PG w Gdańsku,
- Interdyscyplinarne Centrum Modelowania Matematycznego i Komputerowego UW w Warszawie,
- Narodowe Centrum Badań Jądrowych w Świerku (od 2019 roku),
- Poznańskie Centrum Superkomputerowo-Sieciowe IChB PAN w Poznaniu,
- Wrocławskie Centrum Sieciowo-Superkomputerowe PWr we Wrocławiu.

Infrastruktura została wytworzona w serii projektów PLGrid (PL-Grid, PLGrid Plus, PLGrid NG, PLGrid Core). Pierwotnym celem było dostarczenie polskim naukowcom nowoczesnych narzędzi i usług informatycznych w oparciu o zasoby superkomputerowe. W latach 2009-2012 stworzono krajową infrastrukturę gridową, wspierającą badania naukowe realizowane przez zespoły rozproszone geograficznie. Od 2011 roku rozwijano specjalistyczne narzędzia i usługi dla różnych dziedzin nauki.

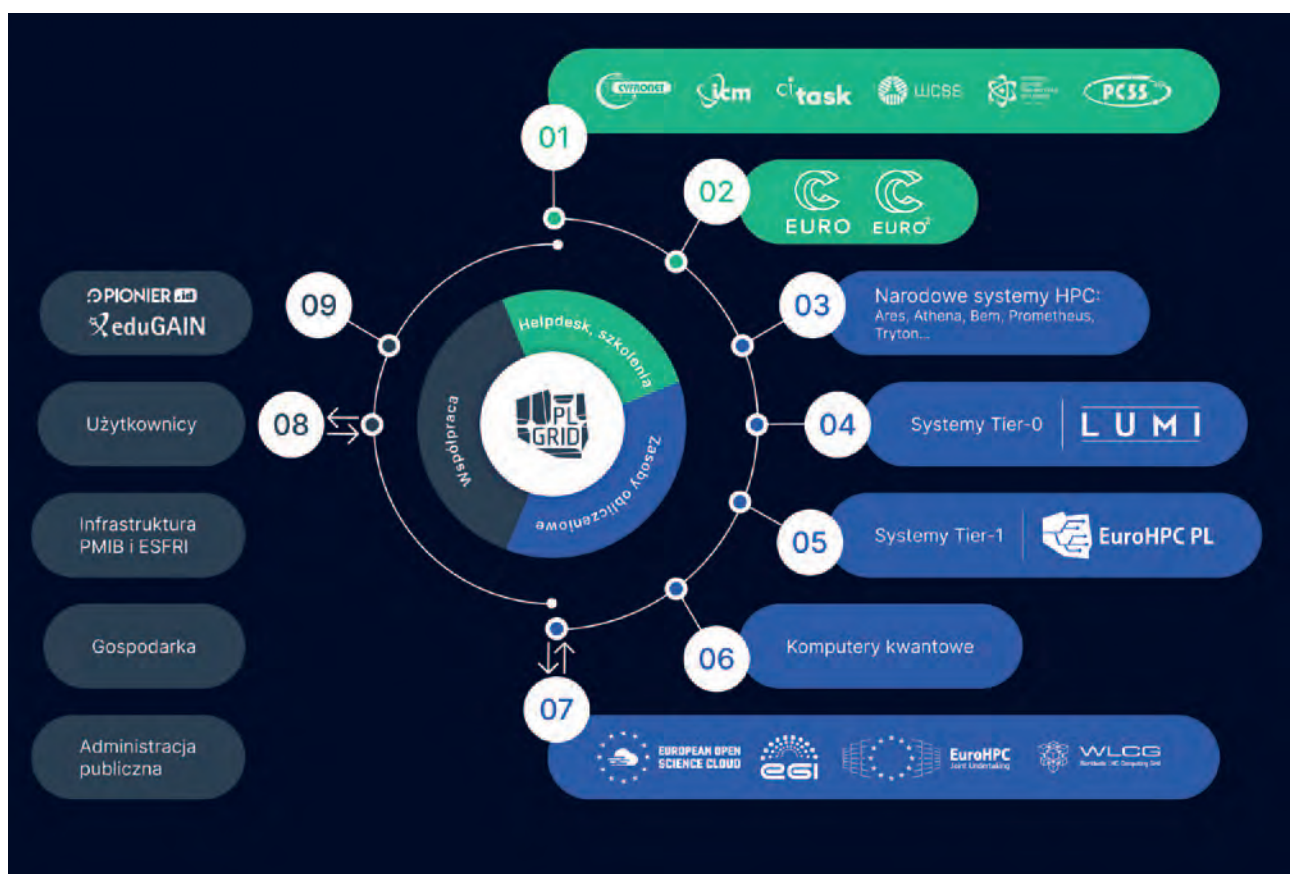
Po rozbudowie narodowej infrastruktury obliczeniowej PLGrid w 2014 roku, możliwej także dzięki jej integracji z infrastrukturą europejską w ramach organizacji dostawców zasobów obliczeniowych i pamięci masowej EGI, użytkownicy PLGrid otrzymali dostęp między innymi do obliczeń w chmurze i nowych narzędzi wspierających obliczenia na wielkich danych. W kolejnych latach wdrożono także nowe, dedykowane usługi obliczeniowe dla grup badawczych z różnych dyscyplin naukowych, uznanych za priorytetowe w Krajowym Programie Badań.

Od 2020 roku w ramach PLGrid prowadzone są aktywności związane z EuroHPC JU, europejskim przedsięwzięciem mającym na celu rozwój eksaskalowej infrastruktury obliczeniowej. Pod egidą EuroHPC JU powstały Narodowe Centra Kompetencji EuroHPC, które umożliwiają dostęp do światowej klasy superkomputerów oraz zapewniają wsparcie technologiczne i szkoleniowe w zakresie wysokowydajnych obliczeń, gromadzenia, przechowywania, przetwarzania i analizowania dużych ilości danych, a także sztucznej inteligencji.

Nieustanny rozwój

W latach 2023-2025, realizując zadania projektu EuroCC 2, prowadzona jest identyfikacja i uzupełnienie deficytów kompetencji w europejskiej infrastrukturze EuroHPC. W Polsce umożliwiono również dostęp do europejskich zasobów HPC użytkownikom z sektora prywatnego i publicznego. Realizowane jest wsparcie we wdrażaniu usług HPC w zakresie sztucznej inteligencji, a od 2023 roku PLGrid oferuje swoim użytkownikom dostęp do platformy pozwalającej na obliczenia z zastosowaniem akceleratorów kwantowych.

W 2024 roku rozpoczęto projekt Narodowa Infrastruktura Chmurowa PLGrid dla EOSC - PLGrid ICON. Projekt ma na celu zbudowanie nowoczesnej infrastruktury chmurowej, umożliwiającej realizację prac badawczo-rozwojowych zgodnie z paradygmatem Otwartej Nauki. PLGrid ICON będzie częścią Europejskiej Chmury dla Otwartej Nauki (EOSC). Chmura PLGrid ICON zapewni środowisko dla prac B+R, ze szczególnym uwzględnieniem zastosowań data science oraz sztucznej inteligencji.



Najszybsze superkomputery w Polsce i Europie

Infrastruktura PLGrid oferuje zarówno dostęp do wszystkich nowoczesnych zasobów HPC należących do ACK Cyfronet AGH (w tym Atheny i Heliosa – najszybszych maszyn w Polsce), jak również do superkomputerów zlokalizowanych w ośrodkach obliczeniowych konsorcjantów. ACK Cyfronet AGH ponadto reprezentuje Polskę w konsorcjum LUMI, w ramach którego, dzięki wkładowi finansowemu państwa i współpracy naszych specjalistów, polscy naukowcy mogą korzystać z zasobów tego najszybszego europejskiego superkomputera do prowadzenia badań bazujących na obliczeniach wielkoskalowych, równoległym przetwarzaniu olbrzymich zbiorów danych i wieloaspektowych analizach, również z wykorzystaniem sztucznej inteligencji.

Superkomputery udostępnione w infrastrukturze PLGrid od lat zdobywają wysokie, prestiżowe miejsca na listach TOP500 – najszybszych superkomputerów i Green500 – najbardziej efektywnych energetycznie superkomputerów. Warto odnotować, że na liście TOP500 z czerwca 2017 roku znalazły się 4 superkomputery PLGrid: Prometheus, Eagle, Tryton, Bem. Natomiast w czerwcu 2024 roku maszyny udostępnione w PLGrid zostały sklasyfikowane na 5 pozycjach: Helios GPU, Lem, Athena, Ares i Helios CPU.

Z kolei na liście Green500 w 2022 roku Athena znalazła się na 10 miejscu, a w 2024 roku Helios GPU zajął wysokie, 3 miejsce.

Oferta i korzyści

Dostęp do Komputerów Dużej Mocy to dla wszystkich osób prowadzących obliczenia naukowe lub na potrzeby gospodarki możliwość wielokrotnego przyspieszenia procesów obliczeniowych. Pozwala to na prowadzenie skomplikowanych symulacji i analiz, także na bardzo dużych zbiorach danych, w znacznie krótszym czasie i zdecydowanie większej skali, co jest obecnie kluczowe w takich dziedzinach jak fizyka, chemia, biologia molekularna czy wytwarzanie nowych materiałów. Dzięki zasobom kwantowym i równoległemu przetwarzaniu ogromnych ilości danych superkomputery wspomagają też rozwój algorytmów AI i ML oraz odkrycia i innowacje w dziedzinie *big data*.

W ofercie PLGrid znajdują się wielkie moce obliczeniowe, zasoby chmurowe i dyskowe oraz dostęp do licznych pakietów oprogramowania naukowego i usług przechowywania danych na światowym poziomie. Dostęp do tych zasobów jest całkowicie bezpłatny dla polskich naukowców i wszystkich osób prowadzących działalność naukową, związaną z polską uczelnią lub ośrodkiem naukowym, w tym studentów i doktorantów.

Infrastruktura to nie tylko najnowocześniejsze maszyny obliczeniowe w Europie, ale także ludzie – i ich ekspercka wiedza. Współpraca z PLGrid zapewnia wszechstronne wsparcie techniczne,

zarówno poprzez ciągły rozwój dokumentacji, jak i codziennie doradzanie użytkownikom PLGrid przez ekspertów w systemie Helpdesk. Prowadzone są również konsultacje i szkolenia informacyjne.

Z myślą o zapewnieniu Użytkownikom wygodnego i intuicyjnego korzystania z zasobów, stworzony został Portal PLGrid (<https://portal.plgrid.pl>), zaprojektowany w oparciu o nowoczesne technologie. Umożliwia on szybkie i bezpieczne wnioskowanie o granty obliczeniowe. Intuicyjny i przyjazny interfejs ułatwia nawigację i korzystanie z naszych usług, a najwyższe standardy zabezpieczeń chronią dane i prywatność Użytkowników.

Codziennym wsparciem w obliczeniach jest również nowa, ujednolicona, bardziej przejrzysta i na bieżąco aktualizowana dokumentacja – PLGrid Guide, dostępna pod adresem <https://guide.plgrid.pl>. Dzięki jej prostej nawigacji znajdą Państwo potrzebne informacje jeszcze szybciej. Dla bardziej zaawansowanych Użytkowników przygotowano również Kompendium wiedzy potrzebnej do tworzenia bardziej wydajnych aplikacji, optymalnie wykorzystujące możliwości obliczeniowe naszych systemów komputerowych (<https://kompendium.plgrid.pl/>).



Przykłady zastosowań

Dostęp do najnowszych technologii znacząco podnosi jakość i efektywność prowadzonych badań, czego rezultatem jest blisko cztery tysiące publikacji powstałych z udziałem PLGrid. Wśród nich można znaleźć szereg ważnych i przełomowych prac z kluczowych obecnie dziedzin wiedzy, w tym tematyki AI, medycyny czy projektów mających na celu przeciwdziałanie skutkom zmian klimatycznych na Ziemi.



Sztuczna inteligencja

Nasi użytkownicy prowadzą liczne badania nad algorytmami sztucznej inteligencji i możliwością ich praktycznego wykorzystania. Przy wsparciu PLGrid powstają modele językowe i prace z zakresu przetwarzania języka naturalnego.

Moc obliczeniowa w nowoczesnej medycynie

Naukowcy w centrum medycyny obliczeniowej Sano (ACK Cyfronet AGH) używają sztucznej inteligencji, zaawansowanych algorytmów bioinformatycznych i symulacji, aby wspierać lekarzy w procesie diagnostycznym i leczniczym. Nowoczesne superkomputery umożliwiają precyzyjne dobranie leczenia najkorzystniejszego dla danego pacjenta i przyspieszają opracowywanie nowych leków, szczepionek i procedur medycznych.

Pogoda i zmiany klimatu

Dzięki superkomputerom polscy naukowcy nie tylko udoskonalają cyfrowe modele pogodowe z myślą o bardziej precyzyjnych i sprawdzalnych prognozach pogody, ale i oceniają globalne zmiany klimatu, a także prowadzą badania mające na celu opracowanie metod ich łagodzenia.

Sięgając gwiazd

Współczesne instrumenty używane w astronomii codziennie skanują niebo, generując przy tym ogromne ilości danych, liczone często w giga i terabajtach. Z drugiej strony modelowanie procesów astrofizycznych wymaga dużych zasobów i zaawansowanych metod obliczeniowych. Polscy astronomowie z powodzeniem wykorzystują PLGrid do zgłębiania tajemnic kosmosu.

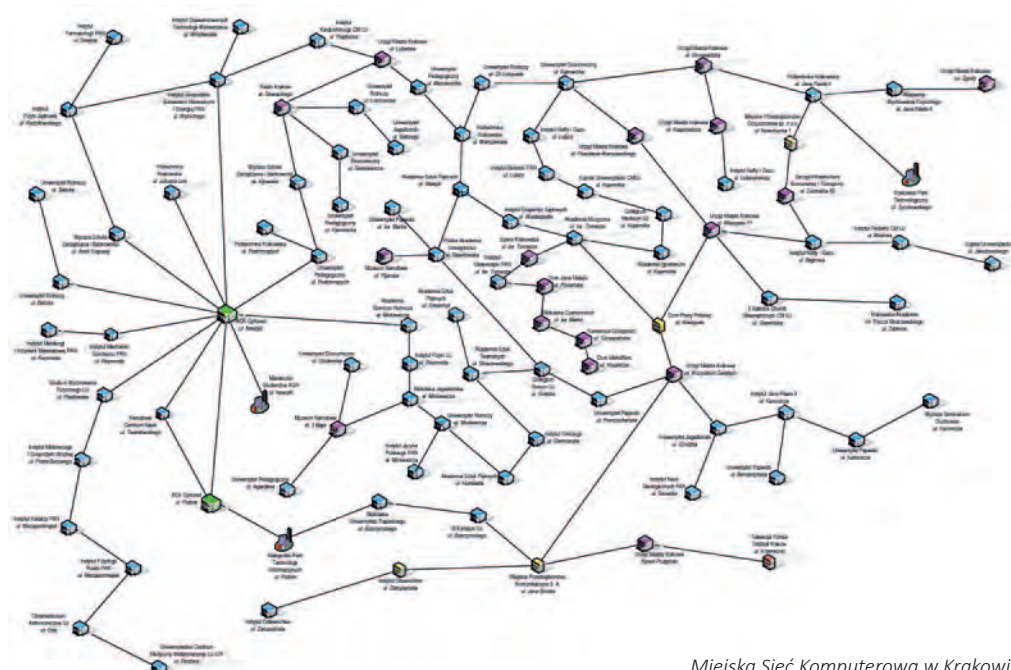
Zyskaj nowe możliwości dzięki PLGrid

Dołączenie do grona Użytkowników PLGrid umożliwia prowadzenie badań na światowym poziomie, z wykorzystaniem najnowocześniejszych narzędzi informatycznych oraz ze wsparciem zespołu ekspertów. Zapraszamy na stronę: <https://www.plgrid.pl/>.



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

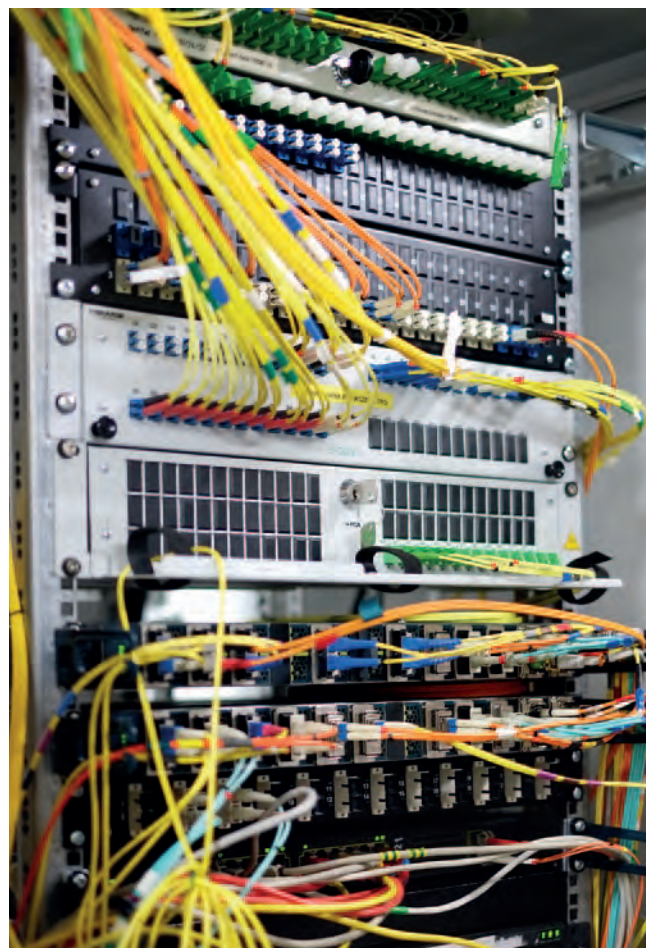
ACK Cyfronet AGH, jako administrator Akademickiej Miejskiej Sieci Komputerowej w Krakowie, stale rozbudowuje i modernizuje sieć, dostosowując ją do potrzeb i oczekiwań użytkowników. Aktualnie Cyfronet eksploatuje własną infrastrukturę światłowodową o łącznej długości ponad 200 km. Światłowody MSK są ułożone w rejonie Starego Miasta, w okolicach kampusu AGH, docierają do Bronowic, Krowdrzy oraz do Czyżyn i Nowej Huty. Zrealizowane zostały także przyłączenia odległych instytutów w Prokocimiu, w rejonie Borku Fałęckiego, w Balicach 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 Akademickiej Miejskiej 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 oraz dosprzętowano podsieć w dwa przełączniki posiadające 8 interfejsów 100 Gb i 16 interfejsów 10 Gb. Wszystkie urządzenia warstwy rdzeniowej są połączone przynajmniej z dwoma (w niektórych przypadkach z 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.

Akademicka Miejska Sieć Komputerowa w Krakowie jest ważnym węzłem akademickiej sieci PIONIER (Polski Internet Optyczny) – ogólnopolskiej szerokopasmowej sieci optycznej stanowiącej istotną bazę dla badań naukowych i prac rozwojowych w Polsce. Sieć akademicka PIONIER jest połączona z Krakowa w kierunku Warszawy, Katowic, Bielska-Białej i Rzeszowa łącami o przepustowości 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.



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 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 Akademickiej Miejskiej 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 Akademickiej Miejskiej 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 2023)

Liczba wysłanych i odebranych wiadomości e-mail	> 19 000 000
Liczba sesji nawiązanych z serwerem pocztowym	> 54 000 000

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 Cyfronet 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. Strona www.cyfronet.pl jest w pełni dostosowana do przeglądania informacji na urządzeniach mobilnych.



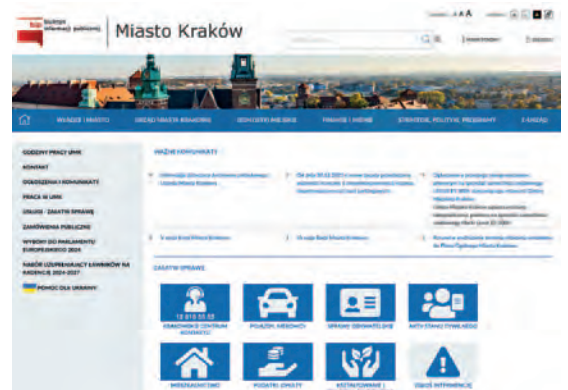
Ważna dla Centrum jest współpraca z Urzędem Miasta. Umowa pomiędzy Gminą Kraków i Cyfronetem, dotycząca promocji miasta, zaowocowała stworzeniem portalu internetowego, który oprócz prezentowania informacji naukowych zapoznaje czytelników z kulturą, zabytkami, turystyką, transportem lokalnym i wieloma innymi aspektami życia w Krakowie.

We współpracy z Urzędem Miasta Krakowa został przygotowany i uruchomiony Biuletyn Informacji Publicznej dla regionu krakowskiego. Na portalu publikowane są Biuletyny Informacji Publicznej Urzędu Miasta Krakowa i kilkuset miejskich instytucji, bibliotek, szkół, przedszkoli, itp.

Od wielu lat ACK Cyfronet AGH aktywnie współpracuje przy tworzeniu oficjalnego serwisu internetowego miasta Krakowa, który funkcjonuje jako Miejska Platforma Internetowa „Magiczny Kraków” (MPI). W MPI udostępnianych jest ponad 40 portali tematycznych, w tym główny portal miejski, dostępny pod adresem www.krakow.pl.



Współpraca z Urzędem Miasta wchodzi również w obszar urządzeń mobilnych – Cyfronet 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. Aplikacja Kraków.pl jest dostępna w kilku wersjach językowych.



Portale MPI i aplikacja internetowa Kraków.pl były wielokrotnie nagradzane i nominowane, np. do World Summit Award jako najlepszy serwis e-administracji w Polsce. Mobilna wersja portalu została natomiast nagrodzona podczas konferencji Mobile Trends jako najlepsza mobilna strona miejska w Polsce.



fot. archiwum AGH



Usługi IT

ACK Cyfronet AGH dysponuje dojrzałą i unikatową infrastrukturą obliczeniową dostępną dla polskich naukowców. Jest ona oparta na kilku 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 Helios, Athena, Ares i Prometheus dostarczają blisko 50 PFlops mocy obliczeniowej.



Przechowywanie danych

150 PB przestrzeni dyskowej i 120 PB zasobów taśmowych na dane wraz z szybkim systemem plików Lustre umożliwia analizę i przechowywanie wielkich zbiorów danych.



Profesjonalne rozwiązania IT

Cyfronet udostępnia szereg zaawansowanych narzędzi umożliwiających łatwy dostęp do rozproszonych zbiorów danych oraz zasobów obliczeniowych, ich analizę, wizualizację oraz udostępnianie.



Centrum danych

Trzy Centra Danych to pełna gwarancja bezpieczeństwa w zakresie udostępniania i utrzymania infrastruktury IT przez cały rok, 24 godziny na dobę, przez 7 dni w tygodniu.

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. 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.

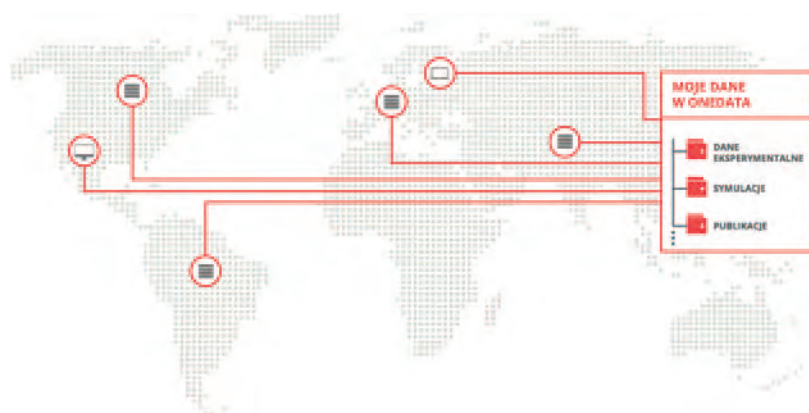
Zaproszenie do współpracy

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.

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 dla 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.

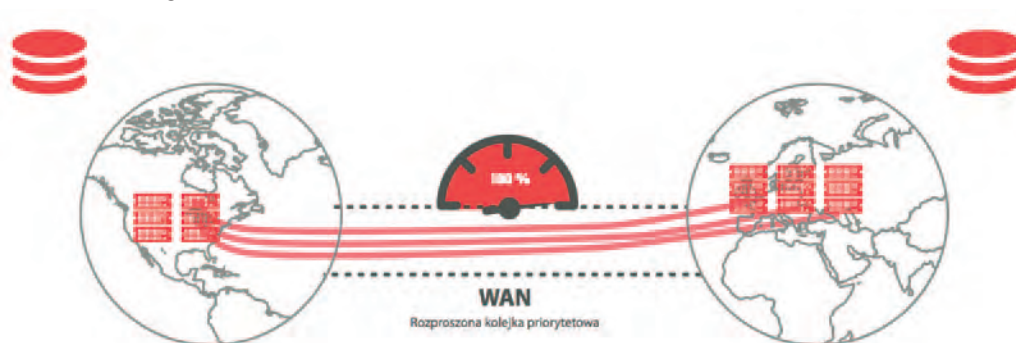
- Funkcjonalność publikowania i dostępu do danych otwartych oraz możliwość łatwej rejestracji identyfikatorów typu DOI dla zbiorów danych.

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 CDMI.
- Elastyczna autentykacja i autoryzacja dostępna w oparciu 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 uruchomiono oraz testowano 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>.

Model Execution Environment

Model Execution Environment (MEE) to stos technologiczny umożliwiający wykonywanie strumieni zadań na infrastrukturach komputerowych wielkiej mocy, w tym na zasobach udostępnianych przez ACK Cyfronet AGH. Celem platformy jest umożliwienie łatwego prowadzenia obliczeń przez naukowców dziedzinowych, tj. przez osoby posiadające podstawową wiedzę z zakresu interakcji z klastrami obliczeniowymi i innymi rodzajami zasobów HPC.



Strumienie, kroki i modele obliczeniowe

W ramach MEE zadania obliczeniowe przyjmują formę tzw. strumieni (*pipelines*), tj. zbiorów obliczeń, z których każde nosi nazwę kroku (*step*), gdzie rezultaty każdego kroku stanowią dane wejściowe dla kolejnych kroków. MEE udostępnia szeroką gamę narzędzi pozwalających użytkownikom na projektowanie kroków, zestawianie ich w strumienie zadań oraz zlecanie wykonywania tych strumieni na dostępnych zasobach obliczeniowych.

Każdy krok jest zdefiniowany w oparciu o zestaw modułów obliczeniowych (tj. wykonywalnego kodu) przechowywanych w repozytorium Git. Moduły te określamy zbiorczo mianem modelu. Gdy pojawia się potrzeba wykonania danego kroku, platforma MEE automatycznie wczytuje dany model na zasoby HPC, a następnie monitoruje jego wykonywanie w oparciu o wybrane dane wejściowe. Przed uruchomieniem strumienia zadań użytkownik jest w stanie wybrać konkretną wersję modelu (dzięki mechanizmom wersjonowania udostępnianym przez Git), co z kolei umożliwia śledzenie rezultatów oraz gwarantuje powtarzalność obliczeń.

Strumienie mogą być uruchamiane w trybie automatycznym lub ręcznym. Strumień automatyczny jest wykonywany w całości, od początku do końca, natomiast w strumieniach ręcznych po każdym kroku następuje przerwa, zaś użytkownik musi ręcznie zlecić wykonanie kolejnego kroku. Mecha-

nizm ten umożliwia weryfikację rezultatów pośrednich i pozwala np. na zaoszczędzenie zasobów poprzez anulowanie strumienia, który nie generuje pożądaných wyników.

MEE jest dostępne poprzez wygodny webowy interfejs UI. Platforma udostępnia również interfejsy programistyczne pozwalające na jej integrację z zewnętrznymi narzędziami.

Zarządzanie danymi badawczymi

MEE zleca i monitoruje wykonanie strumieni zadań na zasobach HPC, a zatem musi też dbać o zarządzanie związanymi z nimi danymi naukowymi, tak aby każdy model operował na właściwych danych i by istniała możliwość pobrania rezultatów z infrastruktury HPC. W tym celu MEE udostępnia zestaw interfejsów zarządzania danymi – użytkownicy mogą przysyłać dane wejściowe i pobierać rezultaty. Platforma zarządza komunikacją z zasobami storage'owymi HPC, umożliwiając automatyczne wprowadzanie danych wejściowych i pobieranie wyników. Opisywana funkcjonalność obejmuje również dedykowane elementy interfejsu użytkownika.

Bezpieczeństwo

Platforma MEE jest zintegrowana z mechanizmami uwierzytelniania i bezpieczeństwa infrastruktury PLGrid. Użytkownicy posiadający konta w infrastrukturze PLGrid mogą uzyskiwać dostęp do MEE za ich pośrednictwem oraz zlecać obliczenia w ramach swoich przydziałów. Każde obliczenie wykonywane jest w ramach konkretnego grantu badawczego, który może zostać skonfigurowany w ramach platformy.

Organizacje

MEE udostępnia poszczególnym zespołom badawczym wyodrębnione środowiska pracy. Każda organizacja posiada unikalny URL dostępowy i może definiować własne strumienie oraz kroki, a także zarządzać własnymi danymi badawczymi. MEE dba o rozgraniczenie organizacji – każdą z nich można zarządzać w autonomiczny sposób, umożliwiając dostęp wybranym grupom użytkowników.

Aplikacje

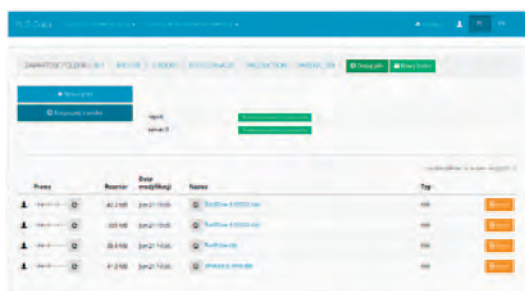
Platforma Model Execution Environment znalazła zastosowanie w szeregu projektów badawczych, do których zaliczamy europejskie projekty EurValve, PRIMAGE oraz InSilicoWorld, konsorcjum POLVAS, Centrum Medycyny Obliczeniowej Sano, a także szereg zespołów powstających na zasadzie ad hoc i dysponujących wyodrębnionymi organizacjami w ramach MEE. Platforma MEE została użyta również jako prototypowa implementacja platformy symulacyjnej wirtualnego bliźniaka tworzonej w ramach europejskiej mapy drogowej przygotowywanej przez projekt EDITH.

<https://mee.cyfronet.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. Obecnie posiada wsparcie dla superkomputerów Ares i Athena, a w przyszłości możliwe będzie rozszerzenie dostępności na inne superkomputery.



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. 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. 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.

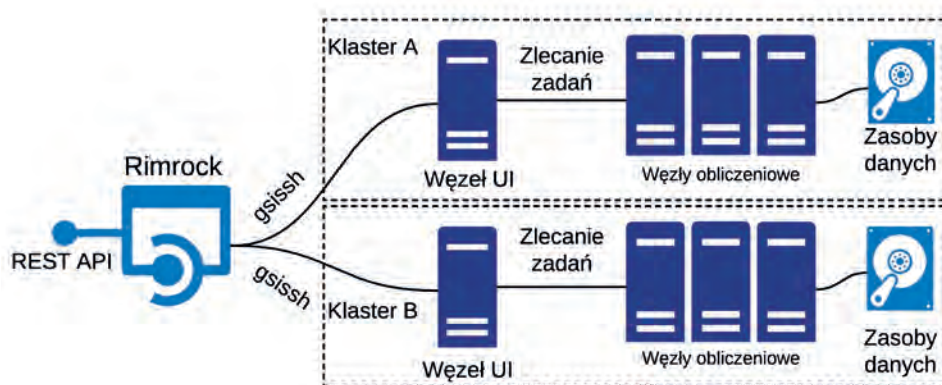
Adres narzędzia PLG-Data: <https://data.plgrid.pl>



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

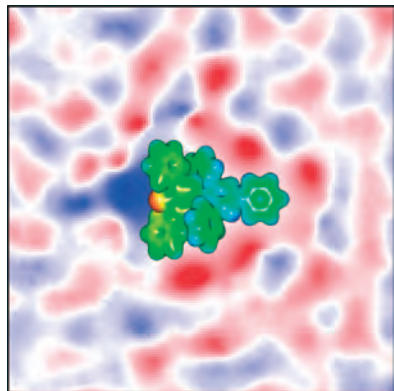
Usługa rimrock wykorzystuje system zarządzania zadaniami Slurm, dzięki czemu zapewnia wsparcie dla jego unikatowych własności. Pozwala to na łatwą integrację aplikacji już istniejących w rozwijanych systemach.

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

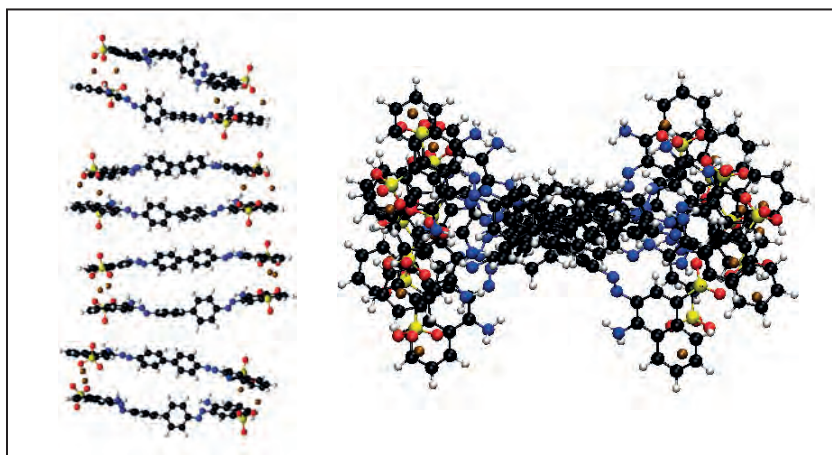
Współczesna chemia obliczeniowa wymaga ciągłego zwiększania zasobów, które umożliwiają przeprowadzenie symulacji dla coraz większych układów (ważnych m.in. w nanotechnologii i naukach biologicznych) oraz poprawę dokładności uzyskanych wyników. Na szczęście, ciągły rozwój technologii w połączeniu z profesjonalnym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

Superkomputery dostępne w naszym Centrum składają się z węzłów obliczeniowych dysponujących nawet 1,5 TB RAM i 48 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:

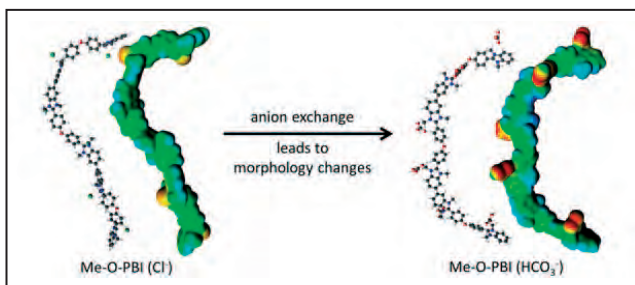
- Wszechstronne i powszechnie używane programy kwantowochemiczne takie jak **Gaussian**, **GAMESS US**, **NWChem**, **Schrödinger**, **Q-Chem**, **Psi4**, **ORCA** czy **TURBOMOLE**, dzięki którym można przeprowadzić obliczenia struktury elektronowej oraz wielu właściwości różnorodnych systemów molekularnych 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 (**AMS** (dawniej **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 **AMS** (**AMSInput**, **AMSVIEW**, 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**.
- Środowisko **AlphaFold** wykorzystujące uczenie maszynowe do analizy struktur geometrycznych białek.
- **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 (per węzeł) wykorzystujących środowisko programistyczne **CUDA**. Wiele pakietów dostępnych w Cyfronecie umożliwia obliczenia GPGPU, wśród nich oprogramowanie chemiczne: **GAMESS**, **Terachem**, **NAMD**, **Quantum ESPRESSO**, **Tinker-HP**. 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 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 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

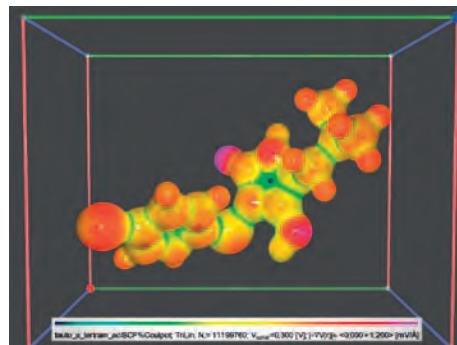
Wizualizacja danych umożliwia analizę i zrozumienie wyników nawet bardzo złożonych obliczeń, szczególnie, jeśli ich forma finalna jest wielowymiarowa bądź uwzględnia czynnik czasu. Większość aplikacji do obliczeń numerycznych posiada jakiś moduł generujący ich wizualizację. W Python'ie jest dostępny 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.

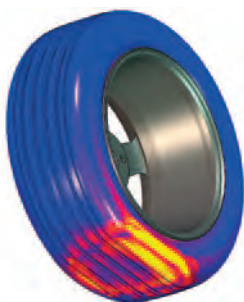
W razie potrzeby developer pakietu służy wsparciem technicznym. Informacje kontaktowe: <https://skos.agh.edu.pl/osoba/janusz-opila-2390.html>.



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

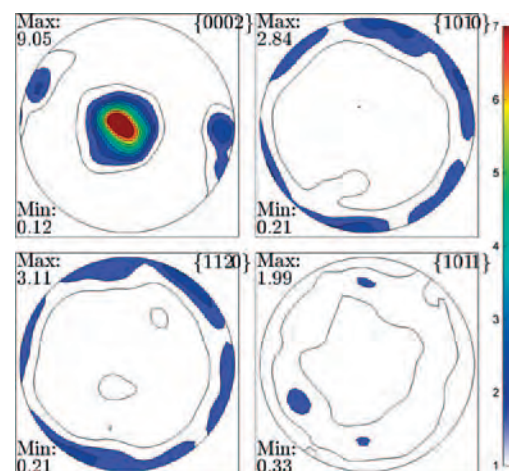
FLUENT umożliwia badanie i symulację zachowania cieczy, co znajduje zastosowanie przy projektowaniu urządzeń wykorzystujących lub transportujących płyny. Może być wykorzystywany w dziedzinie chemii, elektroniki, metalurgii, energetyki cieplnej, biomedycyny i innych.

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, Transformaty 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.

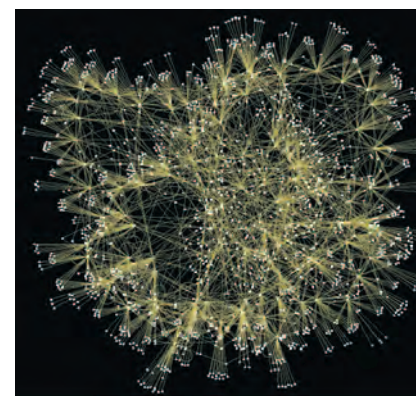


Bartosz Sulkowski: Wyniki symulacji tekstury Zn po wytłaczaniu hydrostatycznym w 250 °C

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 genomowych 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 znajdują 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. Interesującą aplikacją jest również **SageMATH**, pozwalająca na zaawansowane obliczenia matematyczne z użyciem przyjaznego użytkownikowi środowiska Jupyter Notebook. Pozwala ona na użycie wielu dostępnych bibliotek i narzędzi numerycznych, jak NumPy, SciPy, SymPy, Maxima, GAP, Flint oraz R.



Rafał Rak: Sieć połączeń jednodominutowych stóp zwrotu dla spółki KGHM (polska spółka gieldowa)



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

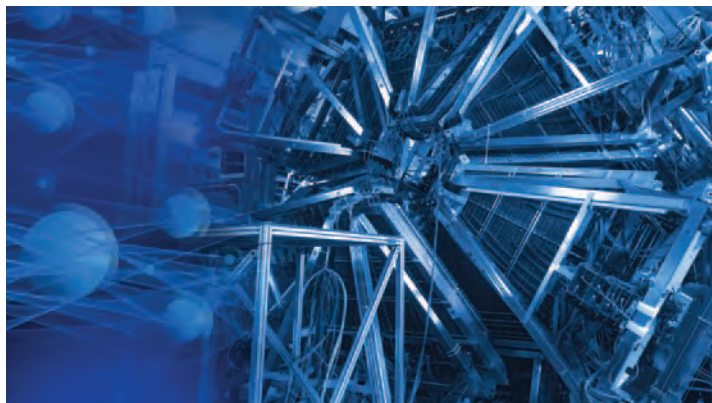
Laboratorium powstało w celu prowadzenia prac badawczych w zakresie wykorzystania komputerów kwantowych w obliczeniach i wsparcia obliczeń klasycznych akceleratorami kwantowymi.

Jednym z kluczowych zadań jest śledzenie rozwoju technologii obliczeń kwantowych oraz dostępnych platform akceleratorów kwantowych celem ich wykorzystania 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 akceleratory kwantowe 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](mailto:m.sterzel[at]cyfronet.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.

Laboratorium nawiązuje współpracę z renomowanymi jednostkami naukowymi i badawczymi w ramach polskich i międzynarodowych projektów. Efektami współpracy są między innymi:

- **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 zuniifikowanego systemu uwierzytelniania użytkownika. Działania EOSC skupiają się wokół realizacji paradygmatu Open Science (otwartej nauki).

W 2022 roku z LPD wyodrębniono Laboratorium Interdyscyplinarnych Obliczeń Naukowych.

Kontakt: Roksana Wilk, r.wilk [at] cyfronet.pl

Laboratorium Interdyscyplinarnych Obliczeń Naukowych

Laboratorium Interdyscyplinarnych Obliczeń Naukowych (LION) prowadzi prace badawcze i rozwojowe dotyczące procesów prowadzenia i wspomagania obliczeń naukowych oraz organizacji danych naukowych. Tworzone przez pracowników Laboratorium portale Science Gateways stwarzają możliwość nawiązywania współpracy z podmiotami zewnętrznymi – polskimi i zagranicznymi.

Zadania LION ukierunkowane są także na:

- realizację grantów naukowych i zleceń przemysłowych,



- wykorzystywanie wiedzy eksperckiej do rozwiązywania problemów wymagających użycia różnych zasobów obliczeniowych, np. technologii uczenia maszynowego,
- popularyzację obliczeń wykorzystujących narzędzia tworzone przez Laboratorium wśród użytkowników,
- wsparcie merytoryczne użytkowników, m.in. poprzez monitorowanie potrzeb, opracowanie dokumentacji, prowadzenie szkoleń.

Zespół LION, współtworząc wcześniej Laboratorium Przetwarzania Danych, 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.

Efektami współpracy LIONu z renomowanymi jednostkami naukowymi i badawczymi w ramach polskich i międzynarodowych projektów jest między innymi rozwój **EPISODES Platform** (<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.

Kontakt: Joanna Kocot, j.kocot [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, ujednolicejący dostęp do danych przechowywanych w rozproszonych systemach. Onedata równie dobrze odpowiada 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 pozwalającemu 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

ACK Cyfronet AGH na Szczycie Cyfrowym IGF Polska we Wrocławiu



4 października 2023 r. przedstawiciele Cyfronetu AGH wystąpili w czasie Internet Governance Forum Polska, gdzie wygłosili prezentację „Superkomputery nie tylko dla nauki – możliwości i dostęp”. Prezentacja i następująca po niej dyskusja obejmowała zastosowania superkomputerów i powiązanych z nimi technologii – sztucznej inteligencji, wielkoskalowego przetwarzania danych i komputerów kwantowych – na rzecz rozwoju nauki, usprawniania funkcjonowania administracji publicznej i wspierania rozwoju przemysłu.

Ponadto zaprezentowana została oferta Narodowego Centrum Kompetencji HPC (EuroCC Polska) i infrastruktury PLGrid w zakresie dostępu do polskich zasobów superkomputerowych, szkoleń i wsparcia użytkowników.

Narodowe Centrum Kompetencji HPC jest koordynowanym przez ACK Cyfronet AGH ośrodkiem eksperckim, a zarazem punktem kontaktowym, który umożliwia użytkownikom korzystanie z polskiej i europejskiej infrastruktury superkomputerowej.



Zespół projektu EuroHPC PL laureatem Polskiej Nagrody Inteligentnego Rozwoju 2023

20 października 2023 r. w czasie Forum Inteligentnego Rozwoju 2023 w Uniejowie ogłoszono laureatów Polskiej Nagrody Inteligentnego Rozwoju. W kategorii „Projekt Przyszłości” nagrodą uhonorowano koordynowany przez Cyfronet AGH projekt „Narodowa Infrastruktura Superkomputerowa dla EuroHPC – EuroHPC PL”. Nagrodę przyznano „za dotychczasowe wyniki przedsięwzięć, mogące skutkować pozytywnym wpływem na rozwój społeczno-gospodarczy, a także otwarte podejście do promocji i komunikacji ze społeczeństwem, w celu przedstawiania wagi korzyści wynikających z realizowanego rozwiązania.”

W imieniu projektu nagrodę odebrał Marek Magryś – ówczesny Zastępca Dyrektora ACK Cyfronet AGH ds. Komputerów Dużej Mocy i koordynator infrastruktury EuroHPC PL. W wydarzeniu wziął także udział Henryk Baniowski, Zastępca Dyrektora ACK Cyfronet AGH ds. Cyberbezpieczeństwa i Innowacyjnej Gospodarki.

Gala rozdania nagród kończyła dwudniowe Forum Inteligentnego Rozwoju, w czasie którego przedstawiciele ACK Cyfronet AGH promowali dokonania projektu EuroHPC PL oraz możliwości skorzystania z nich przez przedstawicieli nauki i innowacyjnej gospodarki.

W czasie wydarzenia, podczas konferencji „Cyfrowa transformacja i inteligentna automatyzacja z wykorzystaniem analityki i sztucznej inteligencji” Marek Magryś wygłosił prezentację pt.: „Narodowa Infrastruktura Superkomputerowa dla EuroHPC – EuroHPC PL” a następnie uczestniczył w panelu dyskusyjnym.

– Wyróżnienie Nagrodą Inteligentnego Rozwoju pokazuje, że rozwój polskiej infrastruktury obliczeniowej jest kluczowy dla dalszego rozwoju i wsparcia innowacyjności całego kraju – mówił Marek Magryś po wydarzeniu. – Serdecznie dziękuję wszystkim, którzy realizują i wspierają działania na rzecz Narodowej Infrastruktury Superkomputerowej dla EuroHPC – EuroHPC PL, w szczególności zespołom Cyfronetu oraz partnerów projektu.



Dzień Otwarty 2023

Akademickie Centrum Komputerowe Cyfronet AGH corocznie organizuje Dzień Otwarty dla środowiska naukowego Krakowa i Małopolski.

Program wydarzenia obejmuje:

- prezentację działalności Cyfronetu i informacje o zasobach sprzętowych, oprogramowaniu i prowadzonych projektach,
- przedstawienie najnowszych kierunków rozwoju architektur komputerowych i sieciowych, w tym również zasobów pozwalających na wykonywanie obliczeń kwantowych,
- wystąpienia dotyczące możliwości obliczeniowych superkomputerów i ich zastosowania do analizy wyników badań, symulacji, wizualizacji i równoległej analizy ogromnych zbiorów danych (big data), w tym na potrzeby sztucznej inteligencji,
- rozstrzygnięcie konkursu na najlepszą pracę doktorską zrealizowaną w oparciu o zasoby obliczeniowe ACK Cyfronet AGH i prezentacje laureatów.

Uczestnicy Dnia Otwartego mają okazję bezpośredniego zapoznania się z unikatowymi usługami Cyfronetu i możliwością ich zastosowania w badaniach na rzecz synergii nauki i gospodarki.





Konferencja projektu EuroHPC PL



W dniach 7-8 grudnia 2023, w Auditorium Maximum UJ w Krakowie, została zorganizowana konferencja podsumowująca projekt „Narodowa Infrastruktura Superkomputerowa dla EuroHPC - EuroHPC PL”, realizowany w latach 2021-2023. Dwudniowy program konferencji obejmował:

- przedstawienie osiągnięć wypracowanych w ramach każdego zadania oraz ogólnych osiągnięć projektu,
- prezentację zasobów sprzętowych (superkomputery i akceleratorzy) oraz niematerialnych

(procedury, oprogramowanie) wytworzonych w projekcie,

- prezentację zasobów kwantowych udostępnionych dzięki projektowi,
- przedstawienie infrastruktury LUMI i LUMI-Q oraz ich integracji z infrastrukturą EuroHPC PL.

W konferencji wzięło udział 130 osób.





Konferencja Użytkowników Komputerów Dużej Mocy (KU KDM) 2024

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 Cyfronetie. 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 infor-

matyki 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 wymianie doświadczeń pomiędzy naukowcami reprezentującymi te same lub róż-

ne dyscypliny oraz zapoznaniu się z nowymi technologiami oraz 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.

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).



<https://www.cyfronet.pl/kukdm24>

Polska Konferencja Open Science 2024



W dniach 10-12 kwietnia 2024 r. została zorganizowana w Krakowie Polska Konferencja Open Science 2024. Konferencja ta jest jedynym wydarzeniem w Polsce poświęconym zagadnieniom związanym z różnymi aspektami Otwartej Nauki, zrzeszającymi wszystkich interesariuszy powiązanych z gromadzeniem, udostępnianiem i zarządzaniem danymi badawczymi, tworzeniem repozytoriów i infrastruktur naukowych, rozwojem inicjatyw związanych z EOSC, udostępnianiem publikacji, artykułów i monografii w otwartym dostępie, a także kwestiom związanym z otwartością w kontekście ewaluacji naukowej.



Jest to wydarzenie tworzone przez społeczność Otwartej Nauki oraz dla tej społeczności i zarazem jedyne, które zaprasza do wspólnej dyskusji i wymiany doświadczeń naukowców, bibliotekarzy, data stewardów, twórców infrastruktur naukowych, dostawców narzędzi, specjalistów IT, wydawców i różnego rodzaju organizacje. W formule naukowej i branżowej dyskusji konferencja jest platformą kształtującą środowisko Open Science w Polsce, pokazującą doświadczenia, trendy i wyzwania, jakie stoją przed inicjatywami Open Science na świecie, w regionie i w Polsce.

Konferencja nawiązuje do tradycji Pomorskiej Konferencji Otwartej Nauki, której w latach 2017-2022 odbyło się sześć edycji.





Konferencja LUMI Day Poland 2024

W dniu 29 lutego 2024 zorganizowana została konferencja online prezentująca superkomputer LUMI, ścieżkę aplikowania o zasoby i możliwości jego zastosowania przez polskich naukowców.

Superkomputer LUMI, zbudowany przez konsorcjum obejmujące 10 państw, w tym Polskę, zajmuje obecnie 8 miejsce na liście TOP500 superkomputerów o największej mocy obliczeniowej na świecie.

Szczegóły dotyczące LUMI znajdują się na stronie: <https://cyfronet.pl/lumi>.

Polscy naukowcy mogą uzyskać dostęp do LUMI za pośrednictwem krajowej infrastruktury PLGrid (<https://portal.plgrid.pl/>)

Wśród tematów poruszanych na konferencji znalazły się między innymi:

- miejsce LUMI w europejskiej inicjatywie EuroHPC JU,
- naukowe wyzwania, na które można odpowiedzieć dzięki LUMI,
- ścieżka aplikowania o zasoby superkomputera LUMI,
- wsparcie dla użytkowników superkomputera LUMI,
- oprogramowanie w środowisku przed-eksaskalowym,
- przykłady wykorzystania zasobów superkomputera LUMI,
- EuroCC - Narodowe Centrum Kompetencji EuroHPC,
- projekt Epicure.



Superkomputer LUMI, fot. Fade Creative

ACK Cyfronet AGH uczestnikiem ISC High Performance 2024 w Hamburgu

ISC High Performance (The Event for High Performance Computing, Machine Learning, Data Analytics & Quantum Computing) to jedno z najważniejszych wydarzeń dla środowiska dostawców i użytkowników rozwiązań związanych z obliczeniami wielkiej skali. Wydarzenie odbyło się w dniach 12-16.05.2024 i tradycyjnie składało się z konferencji obejmującej tematycznie praktycznie wszystkie obszary związane z zastosowaniem superkomputerów oraz targów branżowych, umożliwiających bezpośrednie prezentacje sprzętu, projektów i rozwiązań.

Przedstawiciele Cyfronetu uczestniczyli w konferencji, zapoznając się z najnowocześniejszymi technologiami oraz przykładami ich zastosowania. W czasie konferencji, Marek Magryś, ówczesny Zastępca Dyrektora Cyfronetu ds. Komputerów Dużej Mocy, odebrał certyfikat uzyskania przez superkomputer Helios trzeciego miejsca na prestiżowej liście Green500 najbardziej efektywnych energetycznie superkomputerów świata. Dodatkowo ogłoszono też listę TOP500 najszybszych superkomputerów świata, na której maszyny Cyfronetu zajęły 4 lokaty.

Jednocześnie Cyfronet po raz czwarty wystąpił na ISC jako wystawca, tym razem pod hasłem „Access the future”. Udział Cyfronetu w ISC był m.in. związany z działaniami promocyjnymi w ramach projektu: „Narodowe Centra Kompetencji EuroHPC, faza 2”. Stoisko informacyjno-promocyjne ACK Cyfronet AGH było jednym z nielicznych polskich akcentów w czasie targów. Wśród wystawców znalazły się zarówno firmy – giganci technologiczni, jak Hewlett Packard Enterprise, Intel, czy NVIDIA, jak i czołowe inicjatywy i centra superkomputerowe, w tym m.in. EuroHPC JU, LUMI i Jülich Supercomputing Center. Łącznie w targach uczestniczyło 160 wystawców z 30 krajów.

Główne aktywności Cyfronetu w czasie ISC objęły:

- podtrzymywanie i rozwój relacji z dotychczasowymi partnerami,
- nawiązywanie nowych kontaktów z potencjalnymi dostawcami sprzętu i oprogramowania,
- ocenę nowych rozwiązań technologicznych możliwych do wdrożenia w ramach zarządzanej infrastruktury.

Tegoroczne ISC High Performance przyciągnęło około 3400 uczestników z całego świata.





Cyfronet liderem przedsięwzięć 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.



EuroHPC PL

**Narodowa Infrastruktura Superkomputerowa
dla EuroHPC – EuroHPC PL**



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.

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.

W ramach programu Narodowa Infrastruktura Superkomputerowa dla EuroHPC zrealizowano projekt EuroHPC PL, którego głównym osiągnięciem jest superkomputer Helios.

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.



Narodowa Infrastruktura Chmurowa PLGrid dla EOSC – PLGrid ICON



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, gospodarke 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 ramach programu Narodowa Infrastruktura Chmurowa PLGrid dla EOSC rozpoczęto realizację projektu PLGrid ICON.

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

Bielik - pierwszy polski model językowy

Bielik-11B-v2 powstał w efekcie prac zespołu działającego w ramach Fundacji SpeakLeash oraz Akademickiego Centrum Komputerowego Cyfronet AGH i jest polskim modelem z kategorii LLM (z ang. *Large Language Models*), tj. dużym modelem językowym, posiadającym 11 miliardów parametrów. Prace zespołu nad polskim modelem językowym trwały ponad rok, a ich pierwotny zakres obejmował m.in. zbieranie danych, ich przetwarzanie oraz klasyfikację. Aktualnie zasoby fundacji SpeakLeash są największym, najlepiej opisanym i udokumentowanym zbiorem danych w języku polskim.

Helios i Athena – moc obliczeniowa dla nauki

Superkomputery z Akademickiego Centrum Komputerowego Cyfronet AGH pozwoliły projektowi Bielik rozwinąć skrzydła. Wsparcie naszego zespołu dotyczyło też optymalizacji i skalowania procesów treningowych, prac nad potokami przetwarzania danych oraz rozwoju i działania metod generowania danych syntetycznych, a także prac w zakresie metod testowania modeli. Cenne doświadczenia i wiedza zebrane w wyniku tej współpracy umożliwiły zespołowi ekspertów PLGrid przygotowanie wytycznych oraz zoptymalizowanych rozwiązań, w tym środowisk obliczeniowych do prac z modelami językowymi na bazie superkomputerów Athena i Helios dla potrzeb użytkowników naukowych.

– Zasoby Heliosa, najszybszej aktualnie maszyny w Polsce, wykorzystaliśmy do uczenia modeli językowych – precyzuje Marek Magryś, p. o. Dyrektor ACK Cyfronet AGH. – Nasza rola polega na wsparciu wiedzą ekspercką, doświadczeniem i przede wszystkim mocą obliczeniową procesu katalogowania, zbierania, przetwarzania danych oraz na wspólnym przeprowadzeniu procesu uczenia modeli językowych. Dzięki pracy zespołu SpeakLeash i AGH udało nam się stworzyć Bielika, model LLM, który doskonale radzi sobie z naszym językiem oraz kontekstem kulturowym i który może być kluczowym elementem łańcuchów przetwarzania danych tekstowych dla naszego języka w zastosowaniach naukowych i biznesowych.

Moc obliczeniowa Heliosa i Atheny w tradycyjnych symulacjach komputerowych to łącznie ponad 44 PFlops, a dla obliczeń z zakresu sztucznej inteligencji w niższej precyzji to aż 2 EFlops.

– Jeśli operujemy tak dużymi danymi, jak w przypadku projektu Bielik, to oczywiście infrastruktura potrzebna do pracy przekracza zdolności zwykłego komputera. Musimy dysponować mocą obliczeniową potrzebną do tego, żeby przygotowywać dane, porównywać je ze sobą, trenować modele. Bariera dostępności tego typu superkomputerów powoduje, że mało która firma jest w stanie takie prace prowadzić samodzielnie. Szczęśliwie AGH dysponuje takim zapleczem – wyjaśnia prof. Kazimierz Wiatr.



Dlaczego warto budować polskie modele językowe?

Jedną z korzyści płynących z uruchomienia Bielika jest wzmocnienie pozycji Polski w obszarze innowacji w sektorze AI. Ponadto, co podkreślają twórcy, warto zabiegać o to, aby budować własne narzędzia i tym samym uniezależnić się od zewnętrznych firm, które w przypadku zawirowań na rynkach, regulacji czy ograniczeń prawnych mogą np. uniemożliwić dostęp do swoich zasobów. Tym samym rozwijając i udoskonalając narzędzia w Polsce, budujemy stabilne zaplecze i jesteśmy w stanie zabezpieczać wiele naszych sektorów – bankowy, administracyjny, medyczny lub prawniczy.

Wersja, którą mogą testować użytkownicy, jest utrzymywana nieodpłatnie w domenie publicznej i jest wciąż udoskonalana. Autorzy udostępnili, oprócz pełnych wersji opracowanych modeli, także całą gamę wersji skwantyzowanych w najpopularniejszych dostępnych formatach, które umożliwiają uruchomienie modelu na własnym komputerze.

Przydatne linki:

- Możliwość testowania Bielika: bielik.ai
- Aktualna liczba zebranych danych: speakleash.org/dashboard



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, powódzie 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, Laboratoria Wielkoskalowe, Stanowiska badań geoenergetycznych związane z energią niskowęglową, Tsunami.

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



Źródło: EPOS - European Plate Observing System

wewnątrz i na powierzchni Ziemi. Ma to istotne znaczenie w świetle monitorowania zagrożeń środowiskowych, takich jak: trzęsienia ziemi, powodzie, 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.

Obecnie EPOS wszedł w fazę operacyjną infrastruktury budowanej i wzmacnianej strukturami organizacyjnymi w ramach projektów EPOS PP, EPOS-IP i EPOS SP. Platforma badawcza EPOS Data Portal została udostępniona produkcyjnie w czerwcu 2023 i jest dostępna pod adresem: <https://ics-c.epos-eu.org/>, a portal integrujący informacje o działaniach EPOS pod adresem: <https://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 EPISODES dla jednego z węzłów tematycznych EPOS – węzła Zagrożeń Antropogenicznych (ang. *Thematic Core Service Anthropogenic Hazards - TCS AH*).

Platforma udostępniona pod adresem: <https://EpisodesPlatform.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 udostępniania i wizualizacji danych, Platforma EPISODES zapewnia wirtualną przestrzeń roboczą do uruchamiania aplikacji od różnych dostawców oraz porządkowania danych. Umożliwia także tworzenie własnych aplikacji i wykorzystywanie ich na w/w danych oraz łączenie aplikacji w większe eksperymenty – tworzące tzw. *workflow*. Wspomniana funkcjonalność powstała dzięki krajowemu projektowi POIR – EPOS-PL, natomiast w ramach kolejnego projektu – EPOS-PL+, Platforma EPISODES została wzbogacona o możliwości prowadzenia eksperymentów z wykorzystaniem sztucznej inteligencji.

Oprócz tworzenia i utrzymania Platformy EPISODES, a także dostarczania infrastruktury do obliczeń z jej udziałem, specjaliści z Cyfronetu dbają o utrzymanie elementów integracji pomiędzy węzłem tematycznym TCS AH a portalem centralnym EPOS Data Portal, który umożliwia dostęp do efektów działań we wszystkich węzłach tematycznych EPOS.

W ramach kolejnych projektów europejskich, Platforma EPISODES rozwijana jest także pod kątem zwiększenia ilości dostępnych na platformie danych, adaptacji możliwości platformy dla potrzeb badań przemysłowych oraz integracji serwisów platformy z serwisami European Open Science Cloud (EOSC) – projekty Geo-INQUIRE (Geosphere INfrastructures for QUestions into Integrated REsearch) oraz EPOS ON (EPOS Optimization and Evolution). Platforma rozwijana jest także pod kątem adaptacji na potrzeby badania skrajnych warunków geofizycznych w ramach „bliźniaka cyfrowego” (ang. *digital twin*) budowanego w ramach projektu DT-GEO.





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 za-

awansowanych 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.

Digital Twins

Koncepcja „cyfrowych bliźniaków” (Digital Twins, DT) zakłada wytworzenie wirtualnego modelu jak najwierniej oddającego cechy przyjętego za wzór obiektu. Dzięki temu możliwe jest prowadzenie analiz komputerowych pozwalających podejmować w rzeczywistości efektywne działania, sprawdzone na modelu. Cyfronet poprzez infrastrukturę i wiedzę specjalistyczną wspiera tworzenie modeli „cyfrowego bliźniaka” w różnych dziedzinach nauki.

W ramach projektu EDITH (Ekosystem dla cyfrowych bliźniaków w opiece zdrowotnej) opracowana została architektura zintegrowanego systemu dla realizacji koncepcji „cyfrowego bliźniaka” człowieka, a także wizja jej dalszego rozwoju. Zespół Cyfronetu koordynował prace nad opracowaniem demonstracyjnego prototypu platformy symulacyjnej.

Z kolei projekt DT-GEO (Interdyscyplinarne cyfrowe bliźniaki do modelowania i symulacji złożonych zjawisk w obsłudze środowisk infrastruktury badawczej) zakłada wdrożenie prototypu „cyfrowego bliźniaka” na potrzeby badania skrajnych warunków geofizycznych, jak trzęsienia ziemi (naturalne lub antropogeniczne), osuwiska, wybuchy wulkanów i tsunami. W tym przypadku specjaliści Cyfronetu pracują nad częścią związaną z zagrożeniami antropogenicznymi.

Projekt InterTwin (Interdyscyplinarny cyfrowy silnik dla nauki) współprojektuje i wdraża prototyp interdyscyplinarnego rozwiązania o nazwie Digital Twin Engine (DTE), platformy open source, która zapewnia ogólne i dostosowane komponenty oprogramowania do modelowania i symulacji w celu integracji „cyfrowych bliźniaków” (DT) specyficznych dla aplikacji. Zespół Cyfronetu na rzecz projektu wprowadza integrację z autorską platformą OneData.

Projekt GEMINI (Generacja wieloskalowych cyfrowych bliźniaków pacjentów z udarem niedokrwinnym i krwotocznym) ma na celu dostarczenie zweryfikowanych wielonarządowych i wieloskalowych modeli obliczeniowych do wspomagania decyzji terapeutycznych i lepszego fundamentalnego zrozumienia ostrych udarów, zarówno niedokrwiniowych, jak i krwotocznych. Stworzone zostaną modele wirtualnego bliźniaka, które następnie będą przetestowane przez zespół Cyfronetowy w dużej skali – ten sam model będzie zastosowany dla dużej grupy pacjentów przy użyciu infrastruktury HPC.



Środowisko chmurowe



Możliwość uruchomienia w chmurze kompletnego środowiska pracy, z dostępem do rozproszonych zasobów sprzętowych i pakietów oprogramowania, ułatwia prowadzenie badań w międzynarodowych zespołach. Cyfronet od lat rozwija „ekosystemy” chmurowe, zarówno w oparciu o dostępne narzędzia open source, jak i o autorskie technologie.

Głównym celem projektu INDIGO-DataCloud (INtegrating Distributed data Infrastructures for Global ExpLOitation) było opracowanie środowiska typu PaaS (Platform as a Service) umożliwiającego przeprowadzanie obliczeń dużej skali integrując zasoby gridu oraz cloudu i dostarczając jednolity dostęp zarówno do zasobów obliczeniowych jak i zasobów przechowywania danych. Rolą ACK Cyfronet AGH była m.in. implementacja jednolitego systemu dostępu do danych geograficznie rozproszonych, przechowywanych w heterogenicznych repozytoriach.

Na bazie dokonań projektu przeprowadzono kolejny: XDC (Ekstremalna chmura danych – eXtreme DataCloud), którego celem było zbudowanie specjalistycznych rozwiązań do zarządzania i przetwarzania danymi wielkiej skali w hybrydowej chmurze, wprowadzając tym samym zagadnienia dostępu i migracji danych w rozproszonych środowiskach chmurowych. Cyfronet był odpowiedzialny za realizację systemu rozproszonego zarządzania danymi w środowisku chmurowym.

Jednym z realizowanych obecnie projektów jest DOME (Rozproszony Otwarty Rynek Usług dla Europejskiej Chmury oraz Usług Brzegowych (Edge Services)), w którym w oparciu o założenia programu Gaia-X i otwarte standardy, opracowywane są narzędzia i procedury wspierające rozwój i przyjęcie zaufanych usług Cloud i Edge w Europie. Projekt DOME będzie stanowił punkt dostępu do szeroko rozumianego oprogramowania i usług przetwarzania danych opracowanych w ramach programów UE, takich jak Cyfrowa Europa, Horyzont 2020 lub Horyzont Europa.



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.



Do końca października 2022 roku trwał 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, wspierał proces planistyczny, szukając nowych, skutecznych rozwiązań dla integracji w innych krajach.



W roku 2024 zakończył się projekt EOSC Future (<https://eoscfuture.eu/>), który podwyższył jakość ekosystemu EOSC dla jeszcze lepszego wsparcia europejskiej nauki oraz zachęcił naukowców do korzystania z udostępnianych zasobów. Ten cel został zrealizowany poprzez dalsze prace nad Portalem, integrację kolejnych infrastruktur i inicjatyw naukowych. W ramach projektu Cyfronet odpowiadał za rozwój jednego z trzech głównych komponentów Portalu: części skierowanej do użytkownika, w szczególności za jego elementy wspierające funkcjonalności Marketplace.



Cyfronet uczestniczy także w najnowszych inicjatywach budujących EOSC: FAIRCORE4EOSC, EuroScienceGateway oraz EOSC Beyond. W pierwszym z projektów nasza rola to wsparcie eksperckie w zakresie zarządzania danymi i spersonalizowanego wyszukiwania zasobów EOSC (w tym obiektów naukowych) przy użyciu metod AI. W ramach EuroScienceGateway Cyfronet jest odpowiedzialny za integrację rozwiązań rozproszonej platformy zarządzania danymi Onedata z budowaną infrastrukturą projektu. W EOSC Beyond Cyfronet odpowiada za wsparcie eksperckie oraz dostarczenie rozwiązań w zakresie nowych funkcjonalności EOSC dla naukowców.



EuroHPC – europejska infrastruktura przetwarzania danych

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 34 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).

Pierwszym europejskim eksaskalowym superkomputerem jest JUPITER w Centrum Superkomputerowym Jülich w Niemczech, który w momencie pełnej instalacji w 2024 roku stał się najszybszą maszyną europejską. Drugi superkomputer o mocy przekraczającej moc 1 EFlops stanie w centrum komputerowym TGCC we Francji. Wcześniej jednak uruchamiane były kolejne superkomputery o mniejszej mocy, tworzące infrastrukturę, która pozwala na przeskalowanie technologii i oprogramowania. Do czasu opracowania niniejszej publikacji, EuroHPC JU dostarczyło 3 superkomputery przed-eksaskalowe: LUMI (Finlandia), Leonardo (Włochy), Mare Nostrum 5 (Hiszpania). Dodatkowo pomyślnie uruchomiono operacyjnie 5 superkomputerów petaskalowych: Discoverer (Bułgaria), Karolina (Czechy), MeluXina (Luksemburg), Vega (Słowenia) oraz Deucalion (Portugalia). Kolejnym z tej klasy komputerów będzie Daedalus (Grecja).

W ramach EuroHPC JU finansowany jest również zakup komputerów kwantowych. Do jednego z nich, instalowanego w Czechach, polscy naukowcy uzyskają dostęp za pośrednictwem koordynowanej przez Cyfronet infrastruktury PLGrid.

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



Superkomputer Karolina, Czechy. Fot. EuroHPC JU



Superkomputer Vega, Słowenia. Fot. Atos

Narodowe Centrum Kompetencji HPC

ACK Cyfronet AGH koordynuje działanie polskiego Narodowego Centrum Kompetencji HPC.

Narodowe Centra Kompetencji (NCK) zostały powołane w ramach European High-Performance Computing Joint Undertaking (EuroHPC JU). Wspólne działania mają służyć wsparciu rozwoju innowacyjnych technologii w Europie oraz zwiększyć wykorzystanie infrastruktury HPC (High-Performance Computing) przez użytkowników akademickich, administrację państwową i przedsiębiorców w Europie.

W tym świetle NCK, finansowane poprzez serię projektów EuroCC, mają stanowić centralne punkty kontaktowe w zakresie HPC i technologii powiązanych w swoich krajach. Ich misją jest poprawa i wyrównanie w Europie możliwości wykonywania zaawansowanych obliczeń komputerowych oraz zapewnienie dostępu do infrastruktury i wiedzy specjalistycznej.

EuroCC Polska działa w oparciu o partnerów Konsorcjum PLGrid (ICM, PCSS, WCSS, CI TASK, NCBJ), którego liderem jest ACK Cyfronet AGH. Do głównych zadań Narodowego Centrum Kompetencji należą:

- wspieranie interesariuszy z przemysłu, sektora publicznego i środowisk akademickich we wdrażaniu technologii HPC,
- zwiększenie wykorzystania technologii HPC wśród MŚP w celu podnoszenia ich potencjału innowacyjnego,
- poprawa rozpoznawalności istniejących ofert i usług centrów HPC dla MŚP i przemysłu,
- rozszerzenie portfolio szkoleń,
- stworzenie katalogu usług obejmującego oferty z centrów HPC oraz inne kluczowe usługi i technologie HPC dostawców z Polski,
- podnoszenie świadomości wśród różnych interesariuszy, w tym centrów doskonałości (Centers of Excellence – CoE) oraz europejskich hubów innowacji cyfrowych (eDIH) w Polsce i za granicą.



Przedstawiciele Narodowego Centrum Kompetencji w czasie EuroHPC Summit 2024. Od lewej (w drugim rzędzie): Marek Magryś, Cyfronet; Mateusz Tykierko, WCSS; Mariusz Sterzel, Cyfronet; (w pierwszym rzędzie): Marta Maj (Cyfronet) oraz Szymon Mazurek (Cyfronet)

Szczegóły dotyczące działalności Narodowego Centrum Kompetencji HPC znajdują się na stronie:

<https://cc.eurohpc.pl>.





Superkomputer LUMI

LUMI – nr 3 w Europie, nr 8 na świecie

W listopadzie 2024 ogłoszono kolejną listę TOP500 superkomputerów o najwyższej mocy obliczeniowej świata. Zainstalowany w centrum danych CSC w Finlandii superkomputer LUMI zajął na niej ósme miejsce oraz trzecią pozycję w Europie. Ten niewątpliwý sukces w środowisku HPC (High-Performance Computing) jest efektem współpracy 11 państw, w tym Polski, które wchodzi w skład konsorcjum działającego pod skrzydłami większej inicjatywy, jaką jest EuroHPC Joint Undertaking. Co ważne, dzięki działaniom koordynowanym przez ACK Cyfronet AGH i Ministerstwo, Polacy uzyskali już dostęp do zasobów LUMI.

LUMI – światowa czołówka mocy i wydajności energetycznej

Wspólnie z pozostałymi systemami petaskalowymi oraz przed-eksaskalowymi, w których budowę zaangażowane jest EuroHPC JU, LUMI stanowi element pan-europejskiej infrastruktury, do której zostaną włączone w najbliższych latach również w pełni eksaskalowe komputery. Uzyskanie miejsca w TOP10 superkomputerów to wielki sukces dla państw europejskich, ponieważ pokazuje możliwość konkurowania z systemami z USA, Chin czy Japonii. W testach High Performance Linpack (HPL), superkomputer LUMI wykazał rzeczywistą moc 375 PFlops, a teoretyczną wydajność sięgającą 550 PFlops. Takie osiągi stały się możliwe dzięki architekturze sprzętowej LUMI:

- partycja GPU jest złożona z 2560 węzłów, z których każdy zawiera 64-rdzeniowy procesor AMD Trento i cztery karty AMD MI250X,
- pojedyncza karta MI250X może dostarczyć 42,2 TFlop/s wydajności w testach High Performance Linpack,
- każdy węzeł GPU zawiera cztery karty połączeń sieciowych 200 Gb/s,
- dodatkiem do partycji GPU jest partycja CPU, wykorzystująca 64-rdzeniowe procesory ogólnego przeznaczenia AMD EPYC™ trzeciej generacji.

Jednak przy projektowaniu i budowie infrastruktury LUMI położono mocny nacisk nie tylko na moc obliczeniową, ale też na efektywne wykorzystanie energii. Z tego względu superkomputer LUMI zajął wysokie, 25 miejsce na liście Green500 – superkomputerów o najwyższej wydajności energetycznej, liczonej jako stosunek mocy obliczeniowej do pobieranej energii elektrycznej. Dodatkowo, co warto podkreślić, zasilanie LUMI pochodzi w 100% ze źródeł odnawialnych (elektrownie wodne), a ciepło odpadowe odprowadzane z maszyny jest wykorzystywane do ogrzewania okolicznych budynków. W ten sposób LUMI jest w tej chwili flagowym europejskim przykładem wprowadzania idei zrównoważonego rozwoju do budowy systemów superkomputerowych.

Więcej o superkomputerze LUMI na stronie: <https://lumi-supercomputer.eu>.



Przedstawiciele Cyfronetu w czasie inauguracji LUMI. Od lewej: Mariusz Sterzel, kierownik Laboratorium Obliczeń Kwantowych i Marek Magryś, ówczesny Zastępca Dyrektora ACK Cyfronet AGH ds. Komputerów Dużej Mocy

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 w ramach infrastruktury PLGrid dobrych praktyk.

W okresie od jesieni 2021 roku do wiosny 2022 roku Cyfronet zorganizował dwa konkursy na pilotażowe granty obliczeniowe na superkomputerze LUMI dla polskich naukowców. W okresie testowania infrastruktury zrealizowano 5 polskich projektów. Kolejne konkursy dotyczą już standardowego, rocznego dostępu do infrastruktury. Na ich podstawie, do czasu złożenia niniejszej publikacji, obliczenia na LUMI zostały już zrealizowane lub uruchomione przez 31 polskich projektów. Dostęp jest realizowany za pośrednictwem Portalu PLGrid (<https://portal.plgrid.pl>), a dodatkowe informacje znajdują się na stronie <https://cyfronet.pl/lumi>.



Projekt ma na celu dostarczenie nowych rozwiązań dla otwartej nauki i narzędzi dla nauki w kontekście europejskiej chmury otwartej nauki (EOSC) poprzez dostarczenie nowych funkcji platformy EOSC Core. Planowane usprawnienia mają pozwolić aplikacjom naukowym na znajdowanie, komponowanie i dostęp do zasobów otwartej nauki oraz oferowanie ich jako zintegrowanych rozwiązań.



Głównym celem projektu jest dostarczenie otwartego i w pełni operacyjnego ekosystemu FAIR EOSC. Projekt ma za zadanie rozbudowanie infrastruktury EOSC o komponenty zapewniające wsparcie dla paradygmatu FAIR (Findability, Accessibility, Interoperability, Reusability).



Celem projektu jest zdefiniowanie oraz opracowanie nowego i innowacyjnego programu europejskich studiów magisterskich, skupiających się na HPC. W założeniu program wyposaży studentów w kompetencje i wiedzę dotyczącą HPC, wymaganą przez środowisko akademickie i przemysł.



Poprzez rozwój Portalu EOSC projekt doprowadził do rozszerzenia ekosystemu EOSC, zintegrował istniejące infrastruktury i inicjatywy naukowe oraz zaangażował nowe dziedzinowe i narodowe społeczności naukowe.



W ramach projektu dostarczono transparentny i zunifikowany dostęp do danych z obserwacji Ziemi, głównie satelitarnych, w celu wspierania produkcji żywności oraz przemysłu rolno-spożywczego.



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.



Celem projektu była rozbudowa infrastruktury obliczeniowej umożliwiającej współpracę w obszarze problemów o dużej intensywności łączących duże wolumeny danych przekraczających granice jednego centrum danych.



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



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



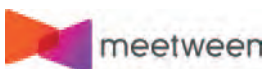
Projekt ma na celu umożliwienie małym i średnim przedsiębiorstwom (MŚP) oraz start-upom w Europie przeprowadzenie eksperymentów z wykorzystaniem infrastruktury obliczeniowej High Performance Computing (HPC) i metod sztucznej inteligencji (AI) w celu poprawy ich konkurencyjności i potencjału innowacyjnego.



Celem projektu jest kontynuacja tworzenia sieci Narodowych Centrów Kompetencji EuroHPC w Europie poprzez współpracę, wymianę najlepszych praktyk i wiedzy oraz podnoszenie krajowego i europejskiego poziomu świadczenia usług HPC. Narodowe Centra Kompetencji świadczą usługi użytkownikom z innowacyjnej gospodarki, środowisk akademickich i administracji publicznej.



Głównym celem projektu jest ustanowienie aktywności wsparcia dla aplikacji użytkowników EuroHPC na poziomach Level 2 i Level 3. W ramach projektu powstanie grupa wsparcia aplikacji, której głównym zadaniem będzie współpraca z biurem oceny wniosków EuroHPC JU i bezpośrednie wsparcie projektów, które uzyskały czas obliczeniowy na zasobach EuroHPC JU.



Projekt ma na celu zbudowanie opartych na nauce rozwiązań technologicznych potrzebnych do zasilania nowej generacji platform videokonferencyjnych dla Europy oraz wspieranie i ułatwianie współpracy biznesowej w całej Unii Europejskiej.



Celem projektu była 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 był polskim etapem rozwoju programu EuroHPC.



Zadaniem projektu było 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.



W projekcie zbudowano unikalne laboratoria badawcze w oparciu o krajową sieć światłowodową PIONIER. Głównym celem projektu było 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 była 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 zostały zwiększone funkcjonalności infrastruktury badawczej EPOS-PL. Powstało nowe Centrum Infrastruktury Badawczej, nowy poligon pomiarowy oraz WNiP: „Platforma IT do badań metodami Sztucznej Inteligencji (EPOS-AI)”.



Zadaniem projektu było 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.



Projekt ma na celu zapewnienie infrastruktury kwantowej dystrybucji klucza oraz sieci komunikacji kwantowej z wykorzystaniem istniejącej infrastruktury światłowodowej sieci PIONIER.



W ramach projektu zbudowane zostało środowisko chmurowe dla treści z obszaru CH (Cultural Heritage) przetwarzające i gromadzące wysokiej jakości treści, dając możliwości uruchomienia wirtualnych przestrzeni dla interdyscyplinarnej i międzysektorowej społeczności współpracującej z obszarem CH.



Projekt zapewni nowy lub znacznie ulepszony dostęp do kluczowych obserwacji, produktów danych i usług, które umożliwiają monitorowanie i symulację dynamicznych procesów w geosferze na niespotykanym dotąd poziomie szczegółowości i precyzji, zarówno przestrzennie, jak i czasowo.



Laureaci poprzednich edycji Konkursu

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, biofizyka, fizyka czy informatyka. 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. 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 edycji 2024



Dr Andrzej J. Kałka Uniwersytet Jagielloński

Rozmowa z autorem pracy:

„Development and Implementation of Computer-Aided Techniques of Data Modeling for Processing and Interpretation of Multicomponent Electronic Spectra Acquired for the Selected Organic Mixtures”

W jaki sposób zainteresował się Pan chemometrią?

Może zabrzmiało to dość szablono, ale głównie za sprawą mojego Promotora. Prawdę mówiąc, już na dość wczesnym etapie studiów zorientowałem się, że zagadnienia typowo kojarzone z chemią, takie jak synteza czy analiza konkretnych próbek, nie wzbudzają zbytnio mojego entuzjazmu. Zdecydowanie bardziej fascynujące wydały mi się (i nadal wydają) zagadnienia związane z odkrywaniem i opisem natury zjawisk, którym te próbki mogą ulegać, i dzięki którym powstały. To przekonanie rzuciło mnie wprost w ramiona chemii fizycznej, teoretycznej i komputerowej. W ścisłym sensie, w arkana chemometrii zostałem wprowadzony w ramach mojej pracy licencjackiej. Chociaż z początku ta tematyka wydawała mi się nieco abstrakcyjna, szybko przekonałem się, że odpowiednie wykorzystanie aparatu matematycznego oraz technologii komputerowej pozwala na wydobywanie z danych, uzyskanych w pomiarach chemicznych, informacji niedostępnych „zwykłym śmiertelnikom”. Urzeczony tą perspektywą, w ramach swojej pracy naukowej postanowiłem poświęcić się rozwojowi tejże metodologii oraz jej praktycznemu zastosowaniu, które – jak już niestety wielokrotnie miałem okazję przekonać – nadal należy do rzadkości.

Co zdecydowało o wyborze pakietu MATLAB jako jednego z głównych wykorzystanych w badaniach środowisk obliczeniowych?

W pewnym sensie nie miałem w tej kwestii większego wyboru – środowisko chemometryczne już na dość wczesnym etapie rozwoju upodobało sobie właśnie ten konkretny pakiet, przez co MATLAB w naturalny sposób stał się (i nadal pozostaje) niejako domyślnym wyborem do prowadzenia obliczeń z zakresu chemometrii. Z perspektywy czasu utwierdzam się w przekonaniu, że wybór ten pozostaje całkiem słuszny. Trzon większości algorytmów stanowią bowiem obliczenia z zakresu algebry liniowej i macierzowej, do których MATLAB (skrót od MATrix LABoratory) został oryginalnie stworzony, tym samym oferując znaczną ilość funkcjonalnych dedykowanych rozwiązań. Poza konsolą, pakiet posiada także wbudowany interfejs graficzny (GUI), co – w porównaniu np. do pakietu R – znacząco ułatwia przetwarzanie oraz wizualizację danych, obejmujących chociażby tworzenie złożonych, wieloelementowych wykresów.

Jakie największe trudności napotkał Pan na drodze badawczej i w jaki sposób stawiał im Pan czoła?

Problemem, który chyba w największym stopniu pokrzyżował moje (i zapewne nie tylko moje) plany, była pandemia SARS-CoV-2, przypadająca na dwa pierwsze lata mojego doktoratu. Praktycznie całkowicie wstrzymana została wówczas krajowa i międzynarodowa wymiana badawcza, co skutecznie uniemożliwiło mi wyjazdy na staże, szkolenia czy konferencje. Na szczęście, wprowadzone ograniczenia nie wpłynęły w znaczący sposób na realizację prowadzonych przeze mnie badań – technologia komputerowa, na której w dużej mierze je oparłem, okazała się odporna na działanie koronawirusa i dobrze odnalazła się w nowej, „zdalnej” rzeczywistości. Drugi problem, o którym według mnie warto jest publicznie wspomnieć, stanowi trudność w pozyskaniu funduszy na prowadzenie własnych badań. Niestety, pula grantów przeznaczonych na sfinansowanie projektów autorstwa młodych naukowców – delikatnie rzecz ujmując – nie należy do największych. Tym samym uzyskanie środków na ich realizację jest często bardzo utrudnione lub wręcz niemożliwe. W moim przypadku bardzo pomocnym okazał się pro-

gram „Inicjatywa Doskonałości – Uniwersytet Badawczy” (IDUB), w ramach którego, w formie tzw. mini-grantów, wsparcie otrzymały projekty zgłoszone właśnie przez doktorantów.

Biorąc pod uwagę zamieszczenie opracowanych algorytmów w bazie RODBUK, jak ocenia Pan zasadność i potrzebę publikowania danych w otwartym dostępie?

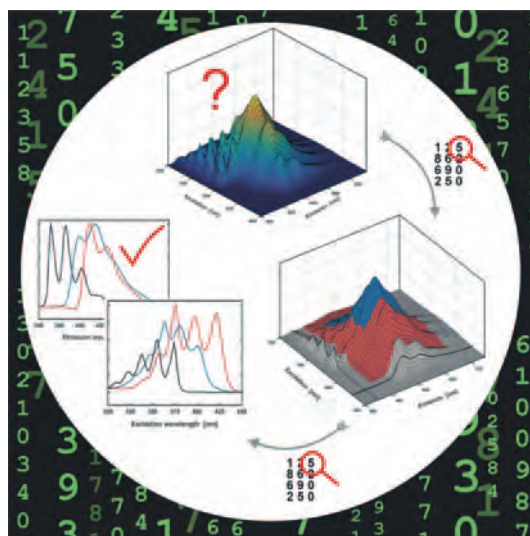
Inicjatywa tzw. „otwartej nauki” z założenia wydaje mi się bardzo słuszna – w mojej opinii wyniki rzetelnie prowadzonych badań w ogólności powinny być powszechnie dostępne dla wszystkich zainteresowanych. Niestety, problem stanowi praktyczna realizacja tego założenia. Publikowanie w formacie „Open Access” (OA) wiąże się z dość dużymi kosztami (rzędu kilku, a nawet kilkunastu tysięcy złotych), na pokrycie których spora część naukowców po prostu nie może sobie pozwolić. W tym aspekcie nie sposób zatem nie pochwalić idei otwartych repozytoriów danych (takich jak np. przywołany RODBUK), w których materiały związane z prowadzonymi badaniami mogą być bezpłatnie zdeponowane, a następnie udostępnione szerokiemu gremium potencjalnych odbiorców.

Na ile zasoby informatyczne udostępniane przez Cyfronet przysłużyły się prowadzonym badaniom?

O ile modelowanie danych eksperymentalnych przy użyciu „zwykłego” komputera nie powinno w praktyce nastroczać większych problemów, o tyle nie mogę sobie wyobrazić prowadzenia z jego pomocą obliczeń kwantowo-chemicznych, stanowiących istotny element mojej pracy badawczej. Dla przykładu, na uzyskanie wyniku jednej z typowych symulacji (których w ramach doktoratu wykonałem dziesiątki, jeśli nie setki) na superkomputerze Ares potrzebowałem nieco ponad dwóch godzin – ta sama symulacja na przeciętnym PC zajęłaby mi szacunkowo pół doby. Wobec powyższego, mogę ze spokojnym sumieniem określić udostępnione zasoby obliczeniowe mianem nieocenionych. Przypuszczalnie moją opinię podzielą również dziesiątki innych użytkowników infrastruktury PLGrid, z którymi – zwłaszcza w ostatnim czasie – dane mi było stawać w szranki o miejsce na podium w systemie kolejkowym slurm.

Co mógłby Pan doradzić początkującym naukowcom, dopiero rozpoczynającym lub planującym rozpoczęcie szkół doktorskich?

Przede wszystkim, aby szczerze przemyśleli swoją motywację i priorytety. Doktorat, szczególnie w ujęciu obowiązującej ustawy (szkoły doktorskie), oferuje sporo możliwości rozwoju naukowego oraz osobistego. Jednakże, aby w pełni z tych możliwości skorzystać, do czego oczywiście gorąco zachęcam, trzeba dysponować odpowiednią ilością czasu, a ten wygospodarować można w zasadzie tylko kosztem ograniczenia lub wręcz rezygnacji z regularnej pracy zarobkowej. To z kolei może odbić się negatywnie na doświadczeniu zawodowym, a co za tym idzie, także na pozycji na rynku pracy (poza sektorem B+R) zaraz po ukończeniu doktoratu. Trzeba zatem szczerze odpowiedzieć sobie na pytanie, czy w perspektywie kolejnych czterech lat ważniejszy jest dla mnie rozwój naukowy, czy zawodowy. Warto w tym miejscu hasłowo wspomnieć o programie tzw. doktoratów wdrożeniowych, oferujących w pewnym sensie opcję pośrednią pomiędzy dwiema powyższymi skrajnościami.



Czym zajmuje się chemometria?
Ideogram zaczerpnięty z A.J. Kałka, A.M. Turek (2021),
J. Fluoresc. 31: 1599-1616



Dr inż. Damian Kułaga

Politechnika Krakowska im. Tadeusza Kościuszki

Rozmowa z autorem pracy:

„Poszukiwanie nowych ligandów receptorów 5-HT_{1A} / 5-HT₇ działających na ośrodkowy układ nerwowy z grupy długołańcuchowych arylopiperazyn i aminotriazyn”

Co sprawiło, że zainteresował się Pan chemią, zarówno w zakresie eksperymentalnym, jak i obliczeniowym?

Moje zainteresowania chemiczne (a w zasadzie life-science) sięgają dość wczesnych lat młodości, kiedy jako nastolatek wykonywałem proste eksperymenty i dzięki temu uczyłem się chemii i poznawałem, jak skonstruowane jest życie na Ziemi. Szkoła średnia i studia coraz bardziej ukierunkowywały mnie w obszarze chemii medycznej, natomiast największy skok mojej wiedzy dokonał się, gdy podjąłem pracę w krakowskiej firmie biotechnologicznej. To właśnie tam poznałem tajniki chemii organicznej oraz możliwości jej zastosowania w projektowaniu leków. Tam też miałem swój pierwszy kontakt z prawdziwym procesem odkrywania leków oraz możliwościami aplikacyjnymi chemii obliczeniowej w tym zakresie. Okazuje się, że nie trzeba syntezować dużych bibliotek związków i następnie sprawdzać ich działania. Można do tego zaprzęgnąć komputer, który pokaże, w którym kierunku chemik powinien podjąć działania syntetyczne, a której grupy związków nie warto syntezować w kontekście określonego działania biologicznego. To znacząco skraca efektywny czas pracy naukowca i ponadto obniża koszty całego projektu naukowego. Mając podstawową wiedzę, postanowiłem ją rozwinąć w ramach własnych badań prowadzonych już w trakcie studiów doktoranckich.

W jakich okolicznościach powstała decyzja, by doktorat poświęcić projektowaniu związków o określonym działaniu farmakologicznym?

Jak już wspomniałem, charakter pracy zawodowej skłonił mnie, a w zasadzie zaszczepił we mnie chęć prowadzenia własnych badań. Chciałem doświadczyć sukcesów naukowych, weryfikując daną hipotezę badawczą, jak i zaznać smak porażki. Szczęśliwy traf sprawił, że moja ówczesna Pani Promotor uzyskała grant naukowy finansowany przez Narodowe Centrum Badań i Rozwoju i w ramach tego grantu pojawił się etat. Oczywiście, nie wahając się, skorzystałem z okazji. Dodatkowym argumentem przemawiającym za tym był fakt, że projekt polegał na opracowaniu związków wykazujących aktywność biologiczną na receptory serotoninowe, wobec czego nadal miałem możliwość doskonalenia warsztatu chemii medycznej, ale również rozwijałem zainteresowania z zakresu chemii obliczeniowej. Wiemy, że choroby centralnego układu nerwowego (takie jak depresja czy ChAD) zaraz po nowotworach i chorobach serca są jedną z najczęstszych przyczyn zgonów. Z tej racji chciałem mieć swój wkład (choćby niewielki) w opracowanie leków czy też środków, które w przyszłości potencjalnie mogłyby stać się lekami pomagającymi osobom chorym.

Dlaczego ważne jest, by poszukiwać ekologicznych metod syntezy nowych związków? Jak ocenia Pan swoje działanie w tym zakresie?

Jest to ważne z kilku powodów. Przede wszystkim nowe metody syntezy pozwalają na lepszą ochronę środowiska, a wiadomym jest, że Unia Europejska mocno stawia na działania w tym zakresie. Kolejny kluczowy powód to oszczędność pieniędzy. Poprzez lepszą optymalizację procesów, redukcję np. czasu syntezy, ilości użytych rozpuszczalników, ilości generowanych odpadów, czy też poprawę ekonomiki atomowej, można zaoszczędzić znaczną sumę środków finansowych. Oprócz tego, pandemia Covid-19 jak w soczewce uwidoczniła, jak wrażliwe są łańcuchy dostaw surowców, których obecnie baza znajduje się głównie w krajach Dalekiego Wschodu lub w Indiach. Możliwość syntezy w rodzimym kraju (lub nawet w krajach UE), ale w oparciu o technologie bezpieczne dla środowiska, pozwoli na uniezależnienie się od krajów azjatyckich, gdzie, jak wiadomo, kwestie ekologiczne traktowane są po macoszemu. Dzięki pracom, które realizowałem w trakcie doktoratu, w znaczny sposób skróciłem czas syntezy związków z nawet kilkudziesięciu godzin do zaledwie 2,5-5 minut. Reakcje prowadzone były w sposób bezrozpuszczalnikowy lub w obecności samej wody. Dzięki temu prowadziłem syntezy w sposób bezpieczny dla środowiska, tani i wydajny, dbając o maksymalną redukcję śladu węglowego oraz wodnego.

W czym przy badaniach pomogły zasoby oferowane przez Cyfronet?

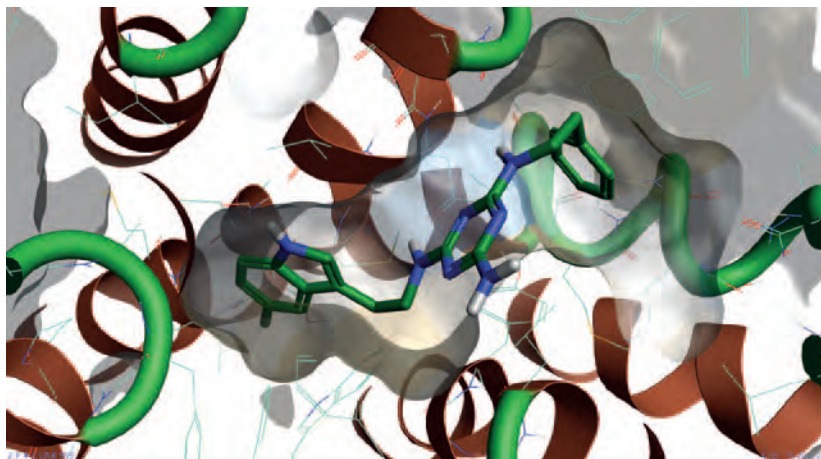
Zasoby te pomogły przede wszystkim w projektowaniu nowych cząsteczek. Korzystając z mocy obliczeniowej superkomputera Prometheus byłem w stanie wykonywać badania związane z dokowaniem oraz dynamiką molekularną dla nowych cząsteczek. Prace te pokazały, które fragmenty cząsteczek odpowiedzialne są za aktywność biologiczną oraz jaką grupę chemiczną należało wybrać do dalszych prac syntetycznych. Badania te nie byłyby możliwe do przeprowadzenia przy zastosowaniu „zwykłego” komputera ze względu na zbyt niskie moce obliczeniowe.

Na co, Pana zdaniem, powinni zwrócić szczególną uwagę naukowcy dopiero wkraczający na ścieżkę prowadzącą do obrony doktoratu?

Na to, żeby byli odporni na przeciwności, których będzie bardzo wiele. Często jest tak, że praca naukowca jest żmudna i nie zawsze prowadzi do oczekiwanych rezultatów. To może demobilizować. Naturalną rzeczą jest, że dany eksperyment nie wyjdzie, ale to też jest dla młodego naukowca wskazówka, co powinien zrobić lepiej, aby prace zakończyły się sukcesem.

Wspomniana żmudność pracy oznacza,

że taka osoba może spędzić w laboratorium wiele godzin oraz jeszcze więcej godzin w domu, na projektowaniu nowych eksperymentów czy szukaniu informacji w literaturze (o ile jest osobą ambitną). Moim zdaniem nie można przesadzać, młody człowiek musi zachować „work-life balance”, gdyż bez tego grozi niestety wypalenie zawodowe, często utrudniające lub nawet uniemożliwiające obronę doktoratu. Jeszcze jedna ważna sprawa, na którą młody naukowiec powinien zwrócić uwagę, to znajomości i nawiązywanie współpracy krajowej i międzynarodowej. Wspólne działanie otwiera możliwości na prowadzenie wartościowych badań interdyscyplinarnych i zawiązywanie konsorcjów naukowych, a także stwarza szanse na potencjalne staże naukowe lub staże typu post-doc.



Zadokowany antagonist receptor 5-HT₁, w miejscu wiążącym



Fot. Jan Zych, PK

Dr inż. Izabela Kurzydym

Politechnika Krakowska im. Tadeusza Kościuszki

Rozmowa z autorką pracy:

„Modelowanie reakcji katalitycznych w procesach deNO_x i deN₂O na drodze Selektywnej Redukcji Katalitycznej”

Co stało u podstaw decyzji, by pracę doktorską poświęcić procesom związanym z redukcją tlenków azotu?

Temat ekologii w nauce jest aktualnie odmieniany chyba przez wszystkie przypadki. Jako mieszkanka Krakowa, na co dzień doświadczam tego, czym jest zjawisko smogu. Chciałam, aby moja praca doktorska, której poświęcę kilka lat, miała wpływ na poprawienie jakości życia. Możliwość zbadania katalizatorów bardziej efektywnych w procesach redukcji tlenków azotu, a więc związków wchodzących w skład zanieczyszczeń powietrza, wydała mi się niesamowicie interesująca. Opracowywanie mechanizmów, odkrywanie zależności, ale i różnic, dało mi ogrom satysfakcji. Od ponad 30 lat Unia Europejska wprowadza coraz bardziej restrykcyjne dyrektywy dotyczące emisji gazów, aby nasze powietrze było czyste a komfort naszego życia lepszy. Dlatego miałam nadzieję, że zakres moich badań może mieć pozytywny wpływ na opracowanie nowych technologii oczyszczania gazów odlotowych w przemyśle, zwłaszcza azotowym, który jest tak istotny z punktu widzenia produkcji m.in. nawozów sztucznych.

Jak mogłaby Pani krótko przybliżyć czytelnikom katalizatory zeolitowe i ich potencjał wdrożeniowy?

Zeolity to struktury krystaliczne składające się z atomów krzemu, tlenu i glinu, czasami z dodatkiem innych atomów metali, jeśli zeolit pochodzi ze źródeł naturalnych. Ze względu na swoją porowatą strukturę posiadają dużą powierzchnię właściwą, co pozwala w łatwy sposób osadzić wewnątrz porów np. metale przejściowe, które z kolei stanowią aktywne centra katalityczne. Dodatkowo zeolity charakteryzują się termostabilnością oraz brakiem toksyczności, co znacząco ułatwia ich wykorzystanie w procesach przemysłowych, często przebiegających w wysokich temperaturach. Ponadto, w przypadku reakcji, w których reagenty są w stanie ciekłym lub gazowym, łatwość separacji katalizatora zeolitowego od produktu ma istotne znaczenie dla przemysłu.

Jak, w świetle własnej pracy, ale też w szerszym kontekście, ocenia Pani zasadność komputerowego modelowania procesów przemysłowych?

Mogę tutaj przytoczyć pewną anegdotę. W swojej pracy prowadziłam obliczenia dla procesu redukcji tlenku azotu z amoniakiem, mogącej mieć miejsce równocześnie z redukcją tlenku azotu z podtlenkiem azotu. Na jednej z konferencji przedstawiany był badany przeze mnie zeolit klinoptylolit, który został wdrożony do przemysłowej instalacji redukcji tlenków azotu w gazach odpadowych w procesie produkcji kwasu azotowego. Badacze wykazali obniżenie stężenia tlenku azotu oraz podtlenku azotu w gazach odpadowych, jednak nie rozumieli tego drugiego zjawiska. Dzięki moim obliczeniom mogłam przedstawić mechanizm, który wpływa na obniżenie stężenia podtlenku azotu i pomóc im zrozumieć procesy zachodzące na katalizatorze. Pokazuje to, jak istotne jest prowadzenie połączonych badań eksperymentalnych i teoretycznych, które mogą wzajemnie się uzupełniać.

Do czego i w jakim zakresie wykorzystała Pani zasoby Cyfronetu?

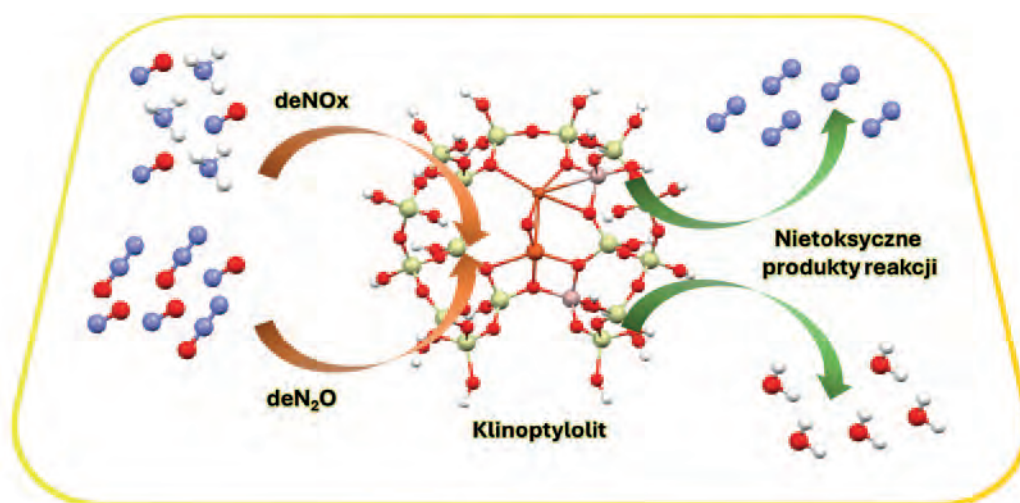
Zasoby superkomputerów Cyfronetu pozwoliły mi na prowadzenie wielu obliczeń jednocześnie. Dzięki dużym mocom obliczeniowym mogłam bez trudu prowadzić obliczenia dla struktur składających się z wielu atomów i otrzymywać wyniki do

analizy w relatywnie krótkim czasie. W swojej pracy doktorskiej zawarłam 72 diagramy opisujące mechanizm deNO_x i deN_2O . Wymagało to wykonania tysięcy pojedynczych obliczeń, co było możliwe tylko i wyłącznie dzięki zasobom Cyfronetu. Miałam również okazję przekonać się, jak prężnie rozwija się infrastruktura superkomputerów w Cyfronecie. W tej chwili kontynuuję swoją karierę naukową korzystając z zasobów nowego superkomputera Ares i z niecierpliwością czekam na możliwość przetestowania Heliosa. Wspaniale obserwuje się rozwój infrastruktury, który przekłada się na zwiększenie efektywności pracy naukowców. Jestem bardzo wdzięczna za możliwość korzystania z zasobów Cyfronetu.

Co mogłaby Pani doradzić badaczom na początku drogi naukowej? Przed czym chciałaby Pani ich przestrzec, do czego zachęcić?

Mam czasem wrażenie, że chemia obliczeniowa budzi wśród młodych naukowców pewien lęk. Wydaje się skomplikowana i mało przystępna. Jednak ważne jest, by nie zrażać się do danego tematu. Dzisiaj chemia i cała nauka rozwija się w stronę wykorzystywania metod obliczeniowych czy sztucznej inteligencji, aby można było dogłębnie poznać świat, który nas otacza. Warto więc chociaż w minimalnym stopniu zawrzeć w swoich pracach chemię obliczeniową, jeżeli nie prowadzoną samodzielnie, to dzięki współpracy z innymi naukowcami.

Młodzi badacze powinni też mieć otwarty umysł i słuchać dobrych rad bardziej doświadczonych naukowców. Jednak warto pamiętać o tym, żeby nie bać się mówić o swoich nowatorskich pomysłach. Dobrze jest wytyczyć sobie główną drogę kariery naukowej, ale należy brać pod uwagę, że odchodzi od niej wiele interesujących ścieżek, na które warto wejść chociaż na chwilę, aby wzbogacić naszą wiedzę o świecie, zarówno tym makro, mikro jak i kwantowym.



Schematyczne przedstawienie połączonego procesu deNO_x i deN_2O na katalizatorze klinoptylolite



Dr Sabina Lichołai

Uniwersytet Jagielloński – Collegium Medicum

Rozmowa z autorką pracy:

„MikroRNA jako potencjalny marker tętniaków aorty brzusznej”

Jak, w sposób niespecjalistyczny, scharakteryzowałaby Pani identyfikację profili molekularnych związanych z tętniakami aorty brzusznej?

Proces identyfikacji profili molekularnych związanych z chorobami naczyń można podzielić na dwa główne etapy: wykonywanie wysokoprzepustowych badań molekularnych oraz przeprowadzanie analiz obliczeniowych. Każdy z tych etapów jest kluczowy dla uzyskania pełnego obrazu molekularnego pacjenta i ma ogromne znaczenie diagnostyczne. Pierwszy etap polega na zastosowaniu nowoczesnych technik biologii molekularnej, takich jak sekwencjonowanie następnej generacji (NGS), proteomika czy metabolomika. Te techniki pozwalają nam na jednoczesne zbadanie tysięcy genów, białek i metabolitów w próbkach od pacjentów. Dzięki tym technologiom jesteśmy w stanie uzyskać szczegółowy profil molekularny, który pokazuje, jakie zmiany zachodzą na poziomie genetycznym, białkowym i metabolicznym w organizmach osób cierpiących na choroby naczyń. Drugi etap to zaawansowane analizy obliczeniowe. Po uzyskaniu surowych danych z badań molekularnych, korzystamy z narzędzi bioinformatycznych i statystycznych, aby zintegrować, przetworzyć i zinterpretować te dane.

Jakie znaczenie mają Pani badania w kontekście diagnostyki i terapii pacjentów z tą chorobą?

Nasze analizy obejmują między innymi porównywanie profili molekularnych pacjentów z grupą kontrolną, co pozwala na identyfikację kluczowych biomarkerów oraz tworzenie modeli predykcyjnych. Modele te mogą pomóc w przewidywaniu przebiegu choroby i odpowiedzi na leczenie. Połączenie dwóch warstw badań, o których wspomniałam wcześniej, pozwala nam na stworzenie kompleksowych profili molekularnych pacjentów, które mogą być użyte w diagnostyce i personalizacji terapii. Dzięki temu jesteśmy też w stanie dokładniej zrozumieć mechanizmy leżące u podstaw różnych zaburzeń, na przykład chorób naczyń, co może w przyszłości prowadzić do opracowania bardziej skutecznych strategii leczenia.

Czy możemy poprosić o komentarz dotyczący potrzeby – ale też możliwości – współczesnego rozwoju diagnostyki w oparciu o narzędzia informatyczne?

Tradycyjnie, diagnostyka molekularna koncentrowała się na analizie białek, które są bezpośrednimi efektorami wielu procesów biologicznych. Metody takie jak ELISA, western blotting, czy spektrometria masowa były powszechnie używane do identyfikacji i ilościowego oznaczania białek w próbkach biologicznych. Chociaż te metody dostarczyły wielu cennych informacji, obecnie poszukiwanie nowych biomarkerów za ich pośrednictwem wydaje się nieco zwalniać, a pula produktów białkowych też jest ograniczona. Nowym, obiecującym kierunkiem w diagnostyce są niekodujące kwasy nukleinowe, takie jak mikroRNA (*miRNA*), długie niekodujące RNA (*lncRNA*) oraz inne małe niekodujące RNA. ncRNA odgrywają kluczową rolę w regulacji genów i są związane z wieloma procesami chorobowymi, co czyni je potencjalnie wartościowymi biomarkerami diagnostycznymi. Identyfikacja i analiza ncRNA wymagają zaawansowanych technologii, takich jak sekwencjonowanie wysokoprzepustowe (NGS – *next generation sequencing*). NGS umożliwia równoczesne sekwencjonowanie milionów fragmentów DNA lub RNA, co jest kluczowe dla odkrywania nowych biomarkerów w dużych i złożonych zestawach danych. Wykorzystanie wy-

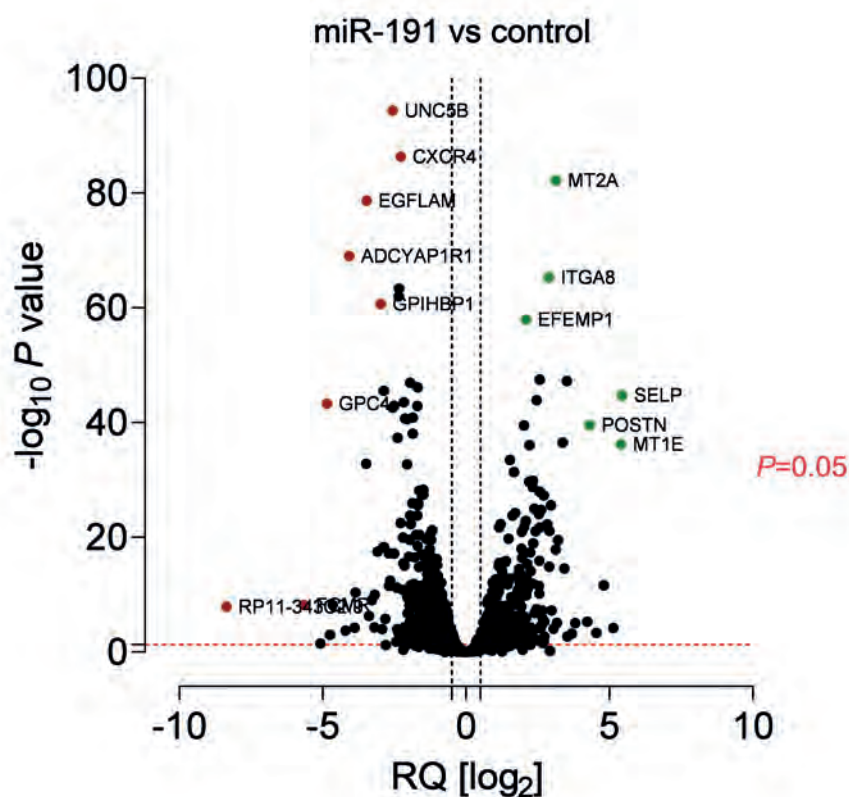
sokoprzepustowego sekwencjonowania generuje natomiast ogromne ilości danych, które muszą być efektywnie analizowane i interpretowane. Tutaj na scenę wchodzi zaawansowane narzędzia informatyczne czy infrastruktura, jak klastry obliczeniowe oraz dedykowane algorytmy bioinformatyczne, które pozwalają na efektywne przetwarzanie tak dużych zbiorów danych.

W jaki sposób, w powyższym zakresie, zasoby Cyfronetu przysłużyły się Pani badaniom?

ACK Cyfronet AGH dostarcza polskim naukowcom możliwość wykonywania badań z wykorzystaniem klastrów obliczeniowych. Zapewniana moc obliczeniowa jest niezbędna do wykonania wymaganych analiz z powodu ogromnych ilości generowanych danych, złożoności analiz oraz potrzeby szybkiego ich przetwarzania. Wykonanie tego typu procedur nie jest możliwe na przykład w oparciu o typowe laptopy czy nawet komputery projektowane do gier. Poza tym, duże ilości danych wymagają nie tylko mocy obliczeniowej, ale także odpowiednich systemów do ich przechowywania i zarządzania. Oferowana przez Cyfronet infrastruktura zapewnia tutaj skalowalne rozwiązania do przechowywania i przetwarzania danych na dużą skalę na potrzeby badań naukowych.

O czym, według Pani, powinni pamiętać studenci i studentki na początku ścieżki doktorskiej?

Podkreśliłabym przede wszystkim wagę odpowiedniego wyboru tematu badawczego. Według mnie powinien być to obszar, który z jednej strony nas mocno interesuje, ale też stanowi swojego rodzaju niszę w dziedzinie. Warto przy tym patrzeć na interdyscyplinarne kierunki badań, które łączą nawet skrajnie różne dziedziny. Takie podejście może prowadzić do nowatorskich rozwiązań, ponieważ umożliwia spojrzenie na problem z różnych perspektyw i wykorzystanie metod oraz narzędzi z różnych dyscyplin. Powiedziałabym też, że warto już na wczesnym etapie uczestniczyć w konferencjach naukowych i spotkaniach branżowych. Pozwala to nawiązać cenne kontakty, wymieniać się doświadczeniami i być na bieżąco z najnowszymi trendami w danej dziedzinie.



Zmiany w zakresie ekspresji genów na poziomie transkryptomu pod wpływem czynników epigenetycznych



Dr Karol Ławniczak

Uniwersytet Łódzki

Rozmowa z autorem pracy:

„Funkcja Wignera na rozmaitościach nietrywialnych topologicznie”

W jaki sposób rozwinęło się Pana zainteresowanie mechaniką kwantową?

Moje zainteresowanie mechaniką kwantową to nic innego jak zainteresowanie tym, jak działa rzeczywistość. To jest to samo zainteresowanie, które napędza dzieci do poznawania świata i zadawania tego irytującego pytania „dlaczego”. Mnie po prostu nie przeszło. Oczywiście, na takie pytania starają się odpowiedzieć nauki przyrodnicze oraz ich filozoficzne „zaplecze”. Fizyka zajmuje się podstawami tego, jak działa rzeczywistość, więc wymaga głębszego namysłu, ale i daje najbardziej satysfakcjonujące wyjaśnienia. Prawdopodobnie to dlatego moim głównym kierunkiem zainteresowań jest właśnie fizyka i to fizyka fundamentalna (choć zajmowałem się też badaniami w innych dziedzinach, m.in. w obszarze ekologii rzek).

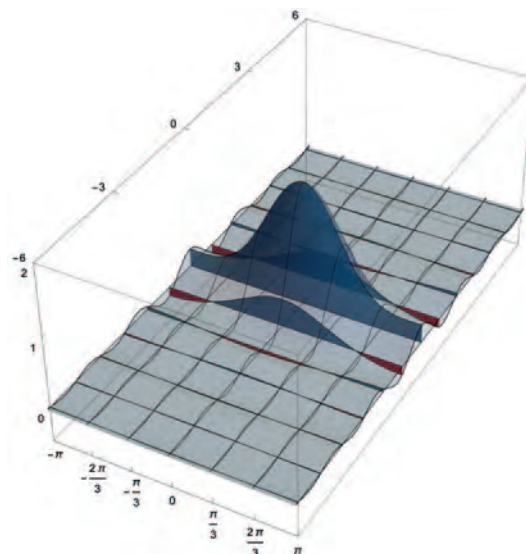
Co zdecydowało o poświęceniu pracy doktorskiej funkcji Wignera?

Czym wyróżnia się Pana podejście w świetle dotychczasowych badań?

W fizyce klasycznej cząstka ma równocześnie dobrze określone położenie i pęd, toteż jej stan jest reprezentowany przez punkt w przestrzeni fazowej. Dla zespołu statystycznego takich cząstek, prawdopodobieństwo znalezienia cząstki w danym obszarze przestrzeni fazowej opisane jest łącznym rozkładem prawdopodobieństwa. W mechanice kwantowej zasada nieoznaczoności Heisenberga wyklucza jednocześnie dokładne określenie położenia i pędu cząstki. Nie jest zatem możliwe sformułowanie rozkładu prawdopodobieństwa na przestrzeni fazowej. Funkcja Wignera jest najbliższym analogiem takiego rozkładu. Sformułowanie zagadnień mechaniki kwantowej w sposób zbliżony do struktury matematycznej mechaniki klasycznej pozwala m.in.:

- dostrzec które z właściwości mechaniki kwantowej różniących ją od klasycznej są istotnymi fizycznie właściwościami tej teorii, a które są artefaktami odmiennego formalizmu matematycznego;
- dogodnie badać zachowanie układów kwantowych przy przechodzeniu do granicy klasycznej;
- nie rezygnować z pewnych intuicji ukształtowanych w wyniku obcowania z fizyką klasyczną.

Definicje obiektów i narzędzi mechaniki kwantowej na przestrzeni fazowej zbudowano przy założeniu, że przestrzeń konfiguracyjna i fazowa układu są topologicznie równoważne przestrzeni euklidesowej. Tymczasem tak naturalne i istotne zagadnienie jak badanie rotacyjnych stopni swobody układów kwantowych, np. w fizyce atomowej i cząsteczkowej, wymaga zaangażowania rozmaitości nietrywialnych topologicznie, gdyż kąty orientacji na płaszczyźnie i w przestrzeni należą odpowiednio do okręgu i sfery – rozmaitości nietrywialnych i to na różne sposoby.



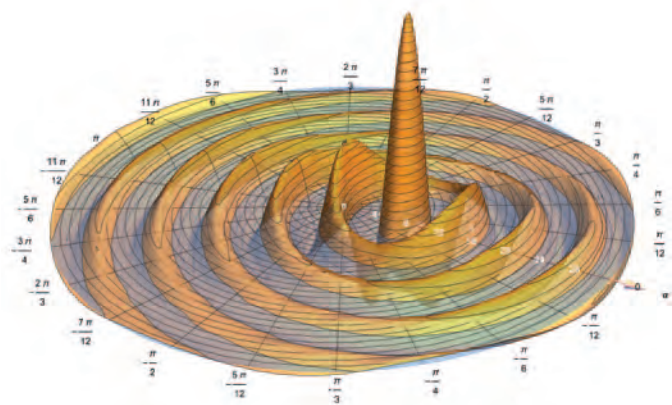
Funkcja Wignera stanu koherentnego na okręgu jako funkcja ciągłej zmiennej kątowej i dyskretnego momentu pędu

Jakie największe wyzwania pojawiły się na Pańskiej drodze badawczej?

Największym wyzwaniem w mojej pracy badawczej były trudności koncepcyjne związane z uogólnieniem obiektów zdefiniowanych w przestrzeni euklidesowej na przypadki topologicznie nietrywialne. Dość powiedzieć, że samo pojęcie średniej musi być przemyślane, gdy uśredniana zmienna należy do dziedziny topologicznie nietrywialnej. Spróbujmy zapisać wzór na średni kierunek wiatru i przetestować go dla trzech punktów danych (azymutów): 1° , 359° , 180° . Już ten przykład wymaga pojęć statystyki na rozmaitościach: średniej zewnętrznej albo, w pewnym sensie lepszej, choć mniej znanej, średniej wewnętrznej. A powiedzieć, że wartości oczekiwane odgrywają istotną rolę w teorii kwantowej to jakby nic nie powiedzieć. Podobnej natury (choć mniej naocznych a bardziej subtelnych) trudności jest więcej i wiele z nich nie miało satysfakcjonującego rozwiązania. Z radością mogę powiedzieć, że część z nich udało mi się rozwikłać i oświetlić na tyle, że możemy je traktować jako wyjaśnione. Kolejnym wyzwaniem były symulacje numeryczne, które okazały się bardziej wymagające niż przypuszczałem u początku pracy.

Jak w tym ostatnim aspekcie wykorzystał Pan zasoby udostępniane przez Cyfronet?

W toku moich badań konieczne było wykonywanie licznych całkowań po kilku zmiennych, w tym po przedziałach nieograniczonych, a także optymalizacji funkcji takie całki obejmujących. Do tego funkcje podcałkowe były szybkozmiennie oraz obejmowały szeroki zakres wartości dodatnich i ujemnych, które kasowały się prawie w całości, ale jednak nie do zera. Przeprowadzenie tych obliczeń tak, aby wynik był wystarczająco precyzyjny, a przede wszystkim wiarygodny, było wymagające obliczeniowo. Na szczęście przynajmniej część z nich można było efektywnie zrównoleglić.



Rozkład quasiprawdopodobieństwa Wignera stanu koherentnego na sferze z niezerowym momentem pędu przedstawiony w przestrzeni momentu pędu

Bez wykorzystania komputera dużej mocy nie dałoby się wykonać ich w rozsądnym czasie. Moje obliczenia prowadziłem na komputerze Ares przy użyciu programu Mathematica. Chciałbym podkreślić, że obliczeń trzeba było wykonać znacznie więcej, niż może sugerować ta ich część, którą przywołano w rozprawie doktorskiej. Taka już jest specyfika pracy badawczej, że zanim odnajdzie się właściwą drogę, próbuje się wielu sposobów, przynajmniej gdy bada się nowy obszar.

Jakie porady mogłyby od Pana otrzymać osoby planujące rozpocząć lub właśnie rozpoczynające ścieżkę doktorską?

Zachęcałbym, aby mniej przejmować się doktoratem, a więcej samymi badaniami. Jeśli mamy ciekawy temat, trzeba go zgłębiać, realizując w nim pasję poznawania rzeczywistości, która, mam nadzieję, przyświeca wszystkim naukowcom czy adeptom Nauki. Z takim podejściem będzie ciekawie, większa też szansa, że uzyskane wyniki będą wartościowe. Wówczas stopień naukowy (gdy już wreszcie spisujemy rozprawę) uzyskamy niejako przy okazji.



Dr inż. Krzysztof Maćkosz

Akademia Górniczo-Hutnicza im. Stanisława Staszica w Krakowie

Rozmowa z autorem pracy:

„Struktura atomowa i transport elektronowy w nanostrukturach kanonicznych izolatorów topologicznych”

Co było dla Pana inspiracją do poświęcenia rozprawy doktorskiej właściwościom nanostruktur chalcogenków bizmutu?

Miniaturyzacja technologii krzemowej to wielki krok naprzód, jednak ciągły rozwój to również napotymane ograniczenia. Połączenie nano-wytwarzania i nowatorskich materiałów to możliwość dalszego rozwoju technologii i krok milowy w dziedzinie komputerów kwantowych.

W jaki sposób otrzymane przez Pana wyniki mogą wpłynąć na rozwój urządzeń bazujących na izolatorach topologicznych?

Świat nanotechnologii to nadal niezbadany teren. Metody wykorzystywane do wytwarzania nano-urządzeń mogą znacząco zmieniać właściwości fizyczne, które są charakterystyczne dla materiałów kwantowych. Dlatego badania nad materiałami w skali atomowej oraz nad metodami wytwarzania w skali nano są dla mnie tak istotne. Pozwalają one na głębsze zrozumienie tych zjawisk i odkrywanie nowych możliwości w dziedzinie nanotechnologii.

Co w pracy naukowej daje Panu największą satysfakcję?

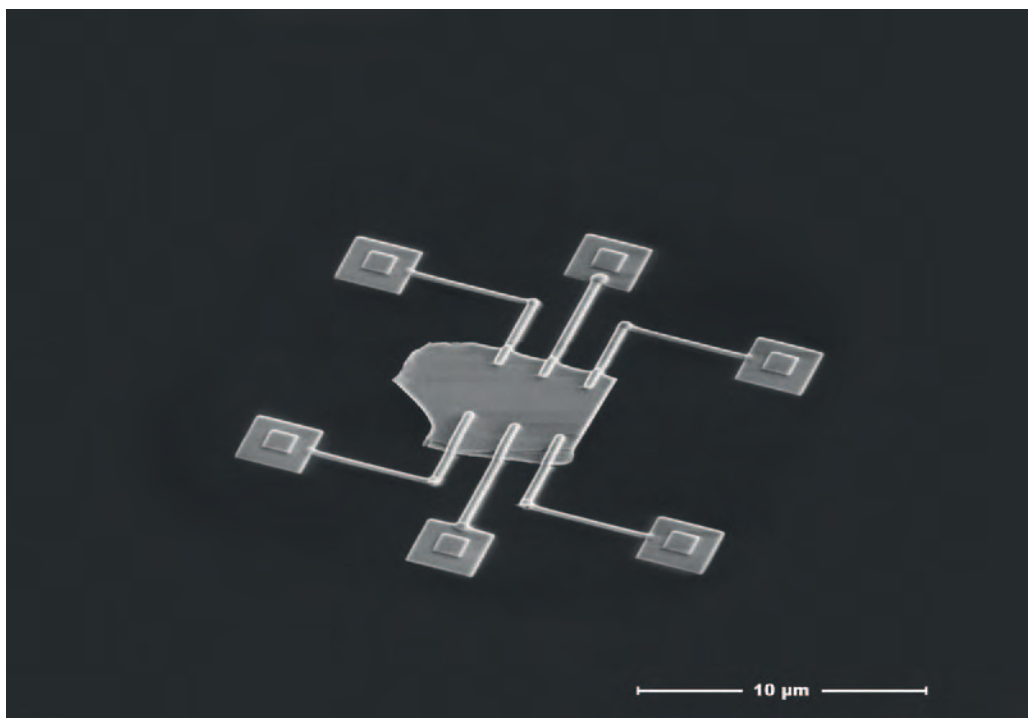
Największą satysfakcję w pracy naukowej daje mi odkrywanie wszechświata i prowadzenie badań, które stanowią mój niewielki kawałek układanki przyczyniającej się do postępu technologicznego. Przede wszystkim jednak cenię sobie możliwość pracy z niezwykle ludźmi i genialnymi naukowcami.

Do czego posłużyły w Pana badaniach zasoby udostępniane przez Cyfronet?

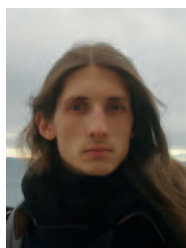
Badania eksperymentalne i teoretyczne to dwie ścieżki w pracy naukowej, które dla wielu nigdy się nie krzyżują. Dla mnie jednak, wykorzystanie zasobów mocy obliczeniowej do badań teoretycznych jest kluczowe, ponieważ pozwala na przewidywanie właściwości fizykochemicznych, co stanowi istotne uzupełnienie moich badań eksperymentalnych.

Gdyby miał Pan udzielić kilku krótkich porad osobom planującym rozpoczęcie lub właśnie rozpoczynającym studia doktoranckie, byłyby to..

Doktorat to długa podróż przez nieznaną, pełną trudności i wyzwań. Trzeba dobrze wiedzieć, do czego się dąży. Osiągnięcie zamierzonego celu przynosi ogromną satysfakcję.



Kryształ izolatora topologicznego, wykorzystywany do badania właściwości transportu elektronowego, uzyskany za pomocą skaningowego mikroskopu elektronowego



Dr Leszek Malec

Uniwersytet Jagielloński

Rozmowa z autorem pracy: „Experimental and *In Silico* Study of Polar Crystals”

Czemu postanowił Pan skupić badania na oddziaływaniach międzycząsteczkowych w kryształach wieloskładnikowych?

Projektowanie nowych materiałów do zastosowania w różnych gałęziach przemysłu oraz medycyny pochłania ogromne ilości zasobów naturalnych, pieniędzy i czasu. Z tego względu, na znaczeniu zyskują metody oparte na przewidywaniu struktury materiału w skali atomowej. Jedną z najważniejszych gałęzi projektowania struktur kryształów wieloskładnikowych w celu uzyskania materiału o pożądanej własności fizycznej jest inżynieria krystaliczna, wykorzystująca zrozumienie oddziaływań międzycząsteczkowych w kontekście upakowania kryształów. Własności kryształów nie są prostą funkcją właściwości poszczególnych elementów struktury. Zamiast tego zależą one silnie od interakcji między cząsteczkami w strukturze oraz ich względnej orientacji i położenia. Zatem, aby w racjonalny i efektywny sposób tworzyć nowe materiały, musimy poznać sposób ich „działania” w skali atomowej w obecnie znanych materiałach funkcjonalnych. Motywacją moich badań było stworzenie metodologii, opartej o obliczenia kwantowo-chemiczne oraz dyfraktometrię rentgenowską, pozwalającej na skorelowanie dynamiki oddziaływań w kryształach z jego makroskopowymi własnościami. Dodatkowo, każdy z trzech badanych typów układów posiadał cechy, których wytłumaczenie umykało dotychczas wykorzystanym metodom eksperymentalnym, co już samo w sobie było ciekawym wyzwaniem badawczym.

Jakie dalsze kroki badawcze mogą być podjęte na podstawie otrzymanych przez Pana wyników?

Zarówno zastosowany schemat badawczy jak i użyte metody i programy mogą zostać zaadaptowane do zbadania wielu własności fizycznych w szerokim spektrum materiałów. Opracowany sposób analizy jest skuteczny w opisie strukturalnych przejść fazowych oraz dynamicznego nieporządku występującego w fazach krystalicznych. Liczę, że zaproponowane przeze mnie podejście, jak również otrzymane wyniki pozwolą lepiej zrozumieć dynamiczne zjawiska strukturalne i związane z nimi efekty makroskopowe, co umożliwi inżynierię nowych materiałów funkcjonalnych. Przykładem mogłyby być materiały ferrokaloryczne, takie jak badany siarczan(VI) amonu, które mogą zostać wykorzystane do stworzenia bardziej efektywnych i przede wszystkim przyjaznych dla środowiska systemów chłodzących. Problem znikania danej fazy krystalicznej jest znany w przemyśle farmaceutycznym, generując ogromne straty w przypadku gdy dana forma polimorficzna zostaje użyta w leku wprowadzonym do produkcji. Uzyskane wyniki, tłumaczące znikanie jednej z odmian polimorficznych kokryształu mocznik-kwas barbiturowy, powinny pozwolić na lepsze zrozumienie tego zjawiska, a także powinny stanowić impuls do dalszych badań nad wpływem porządkowania się cząsteczek w roztworze na efekt krystalizacji.

W pracy wspomina Pan o wykorzystaniu autorskich skryptów do analizy dynamiki układów wiązań wodorowych. Czy możemy poprosić o krótkie przybliżenie oprogramowania?

Głównym celem przygotowanego oprogramowania była analiza korelacji czasowych oraz przestrzennych pomiędzy zmianami parametrów geometrycznych indywiduów chemicznych w badanych układach. Oprogramowanie pozwala na badanie parametrów każdego jonu, cząsteczki oraz wiązania w klastrze obliczeniowym w sposób indywidualny, umożliwiając

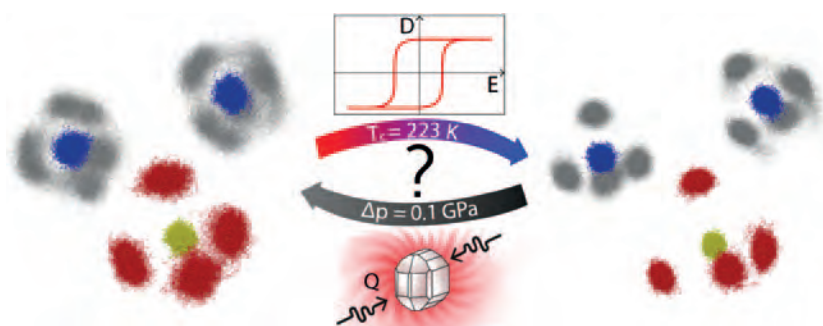
analizę dynamicznego nieporządku w symulowanych strukturach krystalicznych. Możliwe jest także uśrednienie każdego obliczonego parametru w czasie i przestrzeni (z uwzględnieniem symetrii przestrzennej kryształu) pozwalające na porównanie wyników symulacji z danymi uzyskanymi z użyciem dyfraktometrii rentgenowskiej. W programie, każda cząsteczka/ion jest instancją wybranej klasy, która przechowuje współrzędne kartezjańskie wszystkich atomów. Pozwala to na szybkie obliczenie wybranych długości wiązań, kątów walencyjnych i kątów torsijskich dla wybranego typu cząsteczki. Dodatkowo, w niektórych przypadkach zaimplementowałem wyznaczanie bardziej złożonych parametrów geometrycznych, m.in. objętości cząsteczki w oparciu o wyznaczenie jej otoczki wypukłej. Cała część obliczeniowa programu została połączona z modułem graficznym w celu zautomatyzowania wizualizacji wyników.

W jaki sposób wykorzystał Pan zasoby informatyczne udostępniane przez Cyfronet?

Wykorzystując superkomputery Ares oraz Prometheus, wykonałem ponad 70 symulacji metodą dynamiki molekularnej Borna-Oppenheimera dla kilku typów kryształów w programie CP2K. Symulowane z uwzględnieniem okresowych warunków brzegowych klastry zawierały od kilkudziesięciu do kilku tysięcy atomów. Ze względu na wielkość układów i złożoność obliczeń, większość z nich wykorzystywała od 288 do 432 rdzeni. Uzyskane zasoby pozwoliły na analizę wibracyjną każdego układu z wykorzystaniem wygenerowanych widm mocy oraz ich dekompozycji. Otrzymane zasoby dyskowe umożliwiły mi analizę oddziaływań międzycząsteczkowych w badanych kryształach z wykorzystaniem wyżej omówionego, autorskiego oprogramowania. Dodatkowo, dla dwóch badanych układów wykonałem dekompozycję energii oddziaływań dla kilku tysięcy wybranych geometrii z wyznaczonych trajektorii. W tym celu wykorzystałem metodę backfill, wymagającą użycia tysięcy krótkich zadań na klastrze obliczeniowym. Przeprowadzone obliczenia pozwoliły mi wyjaśnić mechanizm przejścia fazowego w ferroelektrycznym siarczanie(VI) amonu, a także opisać zjawisko transferu protonu w trójskładnikowych kryształach zawierających motyw łańcucha polikationowego. Rozwiązałem również zagadkę znikającej formy polimorficznej kokryształu mocznik-kwas barbiturowy, co pozwoliło mi opracować eksperymentalną metodę jej otrzymywania w sposób powtarzalny.

Jaką poradę mógłby Pan przekazać osobom rozpoczynającym szkołę doktorską?

Ze względu na swój unikalny charakter praca naukowa daje dużą swobodę w organizowaniu swojego czasu i tego, czym się zajmujemy. Istotne jest, aby w doborze własnej tematyki kierować się przede wszystkim tym, co nas interesuje i tym, co nasze badania mogą dać innym. Nie należy bać się zmian i przywiązywać się do tego, czym zajmowaliśmy się na wcześniejszych etapach edukacji. Studia doktoranckie dają czas, aby wykształcić swój własny warsztat naukowy, nauczyć się przekazywać swoją wiedzę, a także nawiązać współpracę z innymi naukowcami. Od naszych samodzielnych wyborów zależy, czy nasza praca będzie zarówno satysfakcjonująca jak i wartościowa. Brak radości z tego, co robimy, może łatwo prowadzić do utraty ciekawości, która dla mnie zawsze była główną motywacją do prowadzenia badań. Dlatego od początku doktoratu powinniśmy być świadomi, dlaczego chcemy zajmować się daną tematyką i koncentrować się na tych aspektach badawczych, które pomogą nam się rozwijać.



Graficzna reprezentacja przejścia fazowego w siarczanie(VI) amonu (przedrukowano z L. M. Malec et al. Acta Materialia, 2021, zgodnie z CC BY 4.0)



Dr inż. Agnieszka Winiarska

Instytut Katalizy i Fizykochemii Powierzchni im. Jerzego Habera PAN

Rozmowa z autorką pracy:

„Tungsten aldehyde oxidoreductase from *Aromatoleum aromaticum* – biocatalyst for alcohol production”

Jak wyjaśniłaby Pani przedmiot swoich badań osobie niezwiązanej z dziedziną?

Przedmiotem moich badań było lepsze poznanie enzymu oksydoreduktazy aldehydu z bakterii *Aromatoleum aromaticum* (w skrócie AOR). W swojej pracy pokazałam nową aktywność AOR o potencjalnym znaczeniu praktycznym – redukcję kwasów karboksylowych przy użyciu wodoru cząsteczkowego jako ekologicznego reduktora, pozwalając na konserwację energii np. w formie bioalkoholu. Reakcje te zachodzą w centrum aktywnym zawierającym jon wolframu związany w organometaliczny kompleks. Chcąc poznać mechanizm działania AOR w tej niezwyklej reakcji, musiałam dokładniej zbadać jego strukturę, do czego poza mikroskopią krioelektronową, użyłam również metod modelowania molekularnego.

Czy możliwe jest wyróżnienie przełomowego momentu Pani badań? Jeśli tak, poprosimy o jego przybliżenie?

Charakteryzując AOR odkryłam jego nową, nigdy wcześniej nie odnotowaną dla enzymów wolframowych aktywność hydrogenazy. Oznacza to, że AOR jest zdolne do utleniania wodoru cząsteczkowego rozpuszczonego w wodzie, a uzyskane elektrony wykorzystuje do redukcji kwasów karboksylowych albo mediatorów elektronów takich jak dinukleotyd nikotynoamidoadeninowy (NAD). To odkrycie skierowało moje badania nad AOR na bardziej aplikacyjne aspekty biokatalizy, m.in. na zastosowanie tego enzymu w kaskadach enzymatycznych do syntezy cennych związków.

Ważnym momentem w badaniach było również określenie koordynacji wolframu, dzięki połączeniu badań nad strukturą metodami teoretycznymi z mikroskopią elektronową pojedynczej cząsteczki AOR. Zoptymalizowane metodami klastrową QM i QM:MM geometrie modeli kofaktora o różnych możliwych koordynacjach wolframu zostały zweryfikowane przez porównanie ze strukturą krystaliczną AOR z *P. furiosus* oraz cryo-EM. Badania ujawniły najbardziej prawdopodobną strukturę utlenionego kofaktora wolframowego.

Poprosimy o przedstawienie możliwości aplikacyjnych wyników Pani badań, szczególnie w świetle zgłoszeń patentowych, które ich dotyczyły.

Aktywność AOR jako hydrogenazy może być wykorzystana na konserwację energii np. w formie bioalkoholu oraz w czystej i selektywnej (bez produktów ubocznych) syntezie aldehydów oraz ich pochodnych wykorzystywanych w przemyśle spożywczym, farmaceutycznym i w nowoczesnych materiałach (tzw. hightech). Ze względu na selektywność procesu redukcji kwasu oraz czystą metodę regeneracji NADH, AOR ma duży potencjał do zastosowania jako biokatalizator w nowych

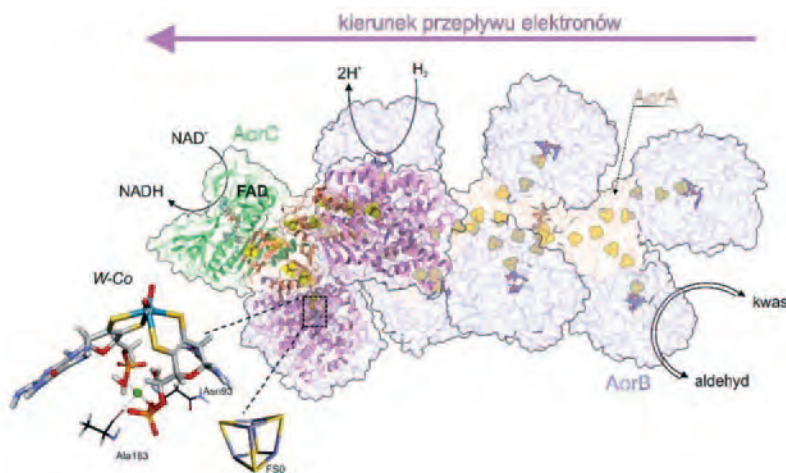
ścieżkach syntezy leków, dodatków do żywności (wanilina) i innych cennych związków. AOR katalizuje również utlenianie aldehydów nawet w bardzo niskich stężeniach, co może być wykorzystane w elektrodach do wykrywania lub usuwania tych często toksycznych związków.

Jak ocenia Pani użyteczność zasobów informatycznych Cyfronetu na potrzeby wsparcia badań naukowych?

Uważam, że zasoby Centrum dają możliwość prowadzenia badań na poziomie konkurującym z wiodącymi ośrodkami tego typu. W swojej pracy skorzystałam z wielu narzędzi zapewnianych przez Cyfronet do modelowania molekularnego metodami QM i QM:MM, do analizy struktury białka i dokowania molekularnego. Możliwość wykorzystania nowych wersji oprogramowania takiego jak Amber czy Gaussian na komputerach dużej mocy pozwoliło mi na zbadanie kofaktora wolframowego w największym, a przez to najdokładniejszym dotychczas optymalizowanym modelu tego typu kofaktora (tj. cały kofaktor zawierający dwa atomy metalu i część organiczną opisany metodami DFT, a enzym metodami mechaniki kwantowej), a muszę podkreślić, że wcześniejsze badania były podejmowane przez wiele grup, m.in. z USA. Bez takich zasobów prawdopodobnie struktura kofaktora dalej byłaby zagadką, a prowadzenie dalszych badań nad mechanizmem reakcji redukcji kwasów karboksylowych niemożliwe.

Co jest, Pani zdaniem, kluczowe przy rozpoczynaniu i na wczesnych etapach ścieżki doktorskiej?

Z pewnością nie jest to zaskakująca odpowiedź, ale uważam, że od początku kluczowy jest networking. Środowisko naukowe jest bardzo otwarte na współpracę i polecam od początku planować staże w innych jednostkach, wspólne projekty i badania. Wiele konkursów wspiera staże i promuje projekty uwzględniające badania w różnych jednostkach. Za takimi działaniami zawsze idą publikacje i kolejne projekty, a po doktoracie ta sieć kontaktów jest szczególnie istotna dla znalezienia pozycji post-doca czy napisania swojego projektu.



Rekonstrukcja filamentowej struktury AOR na podstawie struktury z cryo-EM, fotometrii mas oraz obliczeń QM:MM



Dr Piotr Wróbel

Uniwersytet Jagielloński

Rozmowa z autorem pracy:

„Studies of interactions in solutions based on Molecular Dynamics methods and simulated vibrational spectra”

Co zdecydowało o poświęceniu pracy doktorskiej oddziaływaniom w roztworach wybranych substancji z potencjałem wykorzystania w nowoczesnych bateriach?

Gdy rozpocząłem współpracę ze swoim zespołem w 2017 roku, mój Promotor już realizował grant na prowadzenie badań w tej tematyce. Wtedy dla jednej z takich substancji przeprowadziłem symulacje, które stanowiły treść mojej pracy magisterskiej. Wyniki były zadowalające i wskazywały na zasadność użycia zastosowanej metodologii do tego typu układów. Stąd naturalną była decyzja o kontynuacji działań w tej dziedzinie.

Co w Pańskich badaniach stanowiło największe wyzwanie, jeśli weźmiemy pod uwagę pozyskiwanie danych oraz opracowywanie wyników na ich podstawie?

W kwestii pozyskania danych największym wyzwaniem było otrzymanie sensownego zbioru uczącego dla uczenia sieci neuronowych przewidujących energię potencjalną dla zadanej geometrii układu. Wymagało to starannego doboru różnorodnych geometrii, także takich z „nienaturalnie” porozciągany lub powyginany wiązania, tak aby wyuczony model umiał przewidzieć, że są to struktury niekorzystne energetycznie.

W przypadku opracowania wyników raczej było to liczenie funkcji autokorelacji czasu przebywania, ze względu na złożoność pamięciową algorytmu zastosowanego do jej obliczenia oraz moje – w danym czasie – niewielkie doświadczenie w implementacji tego typu zagadnień.

W pracy przedstawia Pan potencjalne alternatywy dla symulacji za pomocą dynamiki molekularnej ab initio (AIMD). Jak ocenia Pan potrzebę poszukiwania rozwiązań o mniejszym koszcie obliczeniowym?

Uważam, że rozwój metod o mniejszym koszcie obliczeniowym i sensownym kompromisie w kwestii dokładności wyników w stosunku do metod ab initio jest pożądany w kontekście szerszego zastosowania takich badań np. w przemyśle. AIMD, pomimo tego, że prowadzi do rezultatów zbliżonych do wartości eksperymentalnych, wciąż wymaga stosunkowo dużego czasu obliczeniowego – przykładowo otrzymanie 35 ps symulacji dla układu zawierającego 50 cząsteczek węgla etylenu (tzn. 500 atomów) zajęło około miesiąca.

W jaki sposób zasoby Cyfronetu przysłużyły się realizacji części Pana badań związanej z obliczeniami i ich analizą?

Zasoby Cyfronetu właściwie umożliwiły wykonanie jakichkolwiek badań – superkomputer Prometheus a potem Ares były wykorzystywane do prowadzenia wszystkich symulacji (AIMD czy DFTB),

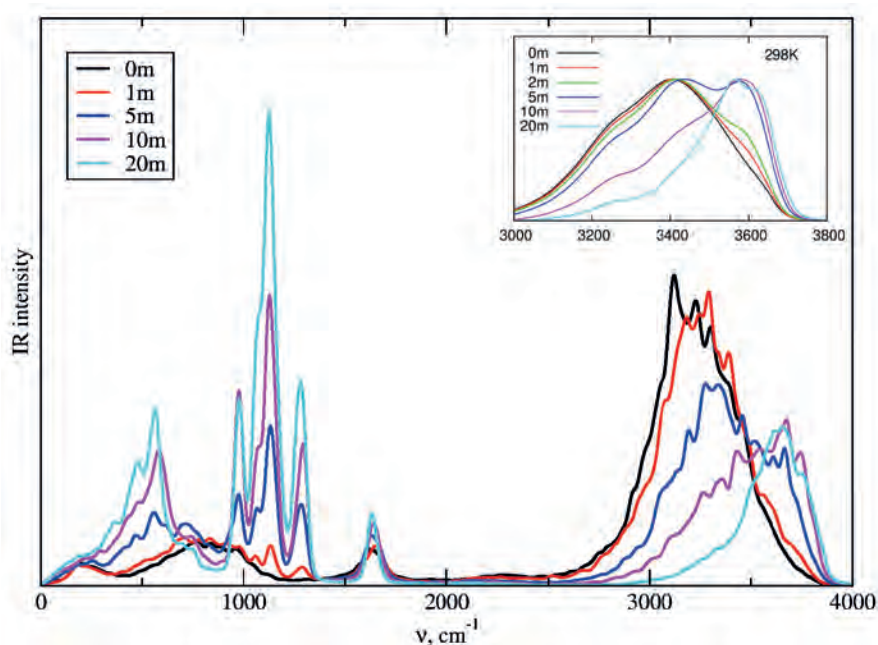
a w późniejszym czasie dedykowana partycja Aresa z układami GPGPU posłużyła do uczenia sieci neuronowych.

Jaki aspekt pracy naukowej daje Panu największą satysfakcję?

Najlepszy jest ten moment zadowolenia, kiedy po długiej pracy, w trakcie której pojawiło się zwątpienie, w końcu zaczyna się klarować wyraźny i oczekiwany wynik.

Co mógłby Pan doradzić osobom, które dopiero rozpoczynają studia doktoranckie? Na co te osoby powinny zwrócić największą uwagę?

Polecam znaleźć swoją niszę, którą się lubi, i nie oglądać się na działalność kolegów z roku. I to mimo tego, że w innym zespole publikuje się więcej, ma się więcej grantów, itd.; nie zapominajmy, że trawa po drugiej stronie ogrodzenia zawsze wydaje się bardziej zielona. Warto korzystać z wyjazdów konferencyjnych w celu oswojenia się z wystąpieniami publicznymi, po tym obrona nie będzie już stresem a formalnością. Dodatkowo ważne jest branie udziału w konkursach na finansowanie badań. Nawet jeśli szanse na wygraną są niewielkie, podstawowa zasada jest prosta – kto nie składa wniosków, ten nic nie wygrywa. Na koniec niestety też bardzo pragmatyczna uwaga – na wypadek, gdyby nie udało się kontynuować kariery naukowej po doktoracie, bezcenne jest zdobycie kwalifikacji umożliwiających inną pracę zawodową. Zapewnia to bardzo duży komfort psychiczny.



Widma w podczerwieni otrzymane z symulacji AIMD dla wodnych roztworów LiTFSI. Wstawka przedstawia eksperymentalne widma w zakresie drgań O-H pochodzące z publikacji: Y. Zhang, N. H. C. Lewis, et. al., *J. Phys. Chem. B* 125 (2021), 4501–4513



Dr inż. Karolina Zawadzińska

Politechnika Krakowska im. Tadeusza Kościuszki

Rozmowa z autorką pracy:

„Synteza nowych nitrofunkcjonalizowanych połączeń heterocyklicznych na drodze [3+2] cykloaddycji”

Jak mogłaby Pani krótko wyjaśnić przedmiot swoich badań osobie niezwiązanej z dziedziną?

Przedmiotem moich badań była synteza i teoretyczny opis biologicznie aktywnych, cyklicznych związków organicznych, które w swojej strukturze zawierają tzw. heteroatomy – tlen i azot. Prościej mówiąc, badałam książkowy przebieg reakcji, które wcześniej udało mi się wykonać eksperymentalnie. Ponadto, w mojej pracy zdeterminowałam potencjalną aktywność biologiczną tych związków, czyli określiłam ich przydatność w zastosowaniu jako leków lub herbicydów.

W badaniach wykorzystała Pani łącznie kilkanaście technik obliczeniowych oraz eksperymentalnych, stanowiących kolejne kroki do osiągnięcia zamierzonych celów. Na ile ta kolejność została z góry ustalona, a na ile wprowadzała Pani elastyczne zmiany w trakcie badań?

W nauce trzeba być elastycznym, ale, rzecz jasna, w granicach rozsądku. Przy opisie teoretycznym reakcji chemicznych zwykle tok obliczeń, jak i zastosowanie danych technik obliczeniowych są założone z góry, jednak w praktyce bywają także odstępstwa od normy. Jako naukowiec, analizując otrzymane wyniki muszę być na tyle elastyczna, aby bazować jedynie na rzetelnych rezultatach i nie naginać teorii tylko po to, by wpasowała się w jakieś założenia. W nauce czasami trzeba dopasować się do badanego obiektu, rzadko bywa tak, żeby to obiekt dopasował się do nas. Szczególnie widać to na przykładzie eksperymentowania w chemii organicznej, gdzie często trzeba dużo kombinować, aby „zmusić” obiekt badań do współpracy.

Jak przedstawia się potencjał aplikacyjny otrzymanych związków?

Obecność heteroatomów w strukturze związków organicznych zwiększa ich biologiczną aktywność, a co za tym idzie, zakres ich potencjalnych aplikacji. Wiele podobnych strukturalnie związków do tych przeze mnie syntezowanych znajduje już zastosowanie jako leki przeciwpadaczkowe, ale przede wszystkim, jako środki przeciwpasożytnicze w weterynarii, jak np. Fluralaner wykorzystywany w popularnych środkach Bravecto dla psów i kotów przeciwko ekto-pasożytom. Te doniesienia literaturowe postanowiłam sprawdzić w odniesieniu do moich związków. Okazało się, że potencjalną aplikacją moich izoksazolin są właśnie środki przeciwbólowe i przeciwpadaczkowe.

Jakie największe przeszkody pojawiły się w Pani pracy badawczej i jak udało się je przezwyciężyć?

Na drodze wykonywania badań do mojej pracy doktorskiej pojawiło się bardzo dużo wybojów i przeszkód. Można powiedzieć, że każdy dzień przynosił nowe wyzwania, z którymi nierzadko ciężko było sobie poradzić. Najtrudniejszym był jednak etap syntezy i oczyszczania otrzymanych

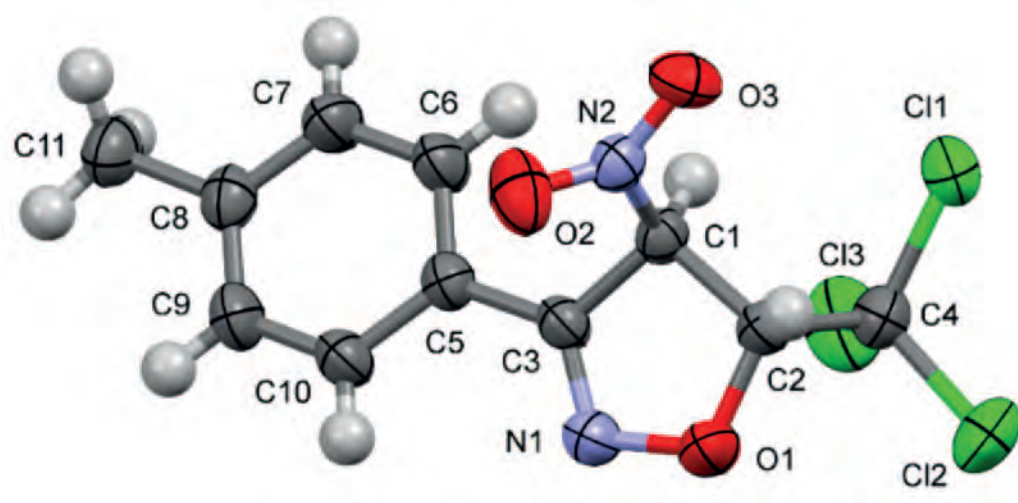
przeze mnie związków. Etap ten wymagał ogromnego poświęcenia i cierpliwości, a także pracy zespołowej.

W jaki sposób zasoby Cyfronetu przysłużyły się do realizacji części badań związanej z obliczeniami i ich analizą?

Bez możliwości korzystania z mocy obliczeniowej superkomputerów Cyfronetu nie mogłabym tak sprawnie przeprowadzić w zasadzie żadnych badań. Wszystkie swoje zaplanowane kroki w pracy badawczej opierałam głównie na modelowaniu komputerowym przebiegu reakcji, żeby wiedzieć, czy jest ona w ogóle możliwa do zrealizowania. Niewątpliwie, obliczenia przyczyniły się do skrócenia czasu poszukiwania błędów i rozwiązań eksperymentalnych, dzięki obliczeniom mogłam zaoszczędzić odczynniki, a wreszcie mogłam skupić się na tych reakcjach, które faktycznie miały jakiś sens.

Jakimi radami lub przemyśleniami mogłaby Pani podzielić się z naukowcami dopiero zaczynającymi szkoły doktorskie?

Na pewno nie poddawajcie się! Prowadzenie badań w różnych dziedzinach, bo nie tylko w chemii, jest bardzo ciężkie, a wiele z nich może skończyć się fiaskiem. Jednak pamiętajcie, że każdy uzyskany wynik jest wartościowy! Bądźcie ciekawi wszystkich szczegółów, bo czasami to one odgrywają główną rolę. Chętnie wychodźcie poza utarte schematy.



Wizualizacja struktury molekularnej izoksazoliny na podstawie pomiarów XRD dyfrakcji rentgenowskiej



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 modułem 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 przepustowości 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 80 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 przepustowości 2x10 Gbps



SGI Origin2000



SGI ALTIX 3700



HP Cluster Platform 3000 BL



Medal jubileuszowy



Hala Maszyn

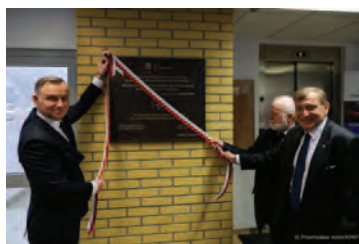


Superkomputer Prometheus

- 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 przepustowości 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 czerwcową) i 59 miejsce (edycja listopadową) na liście TOP500
- 2017** Prometheus zajmuje 71 miejsce (edycja czerwcową) i 77 miejsce (edycja listopadową) 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 czerwcową) i 131 miejsce (edycja listopadową) 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, 2,7 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 4 PFlops.
- Prometheus zajmuje 373 i 440 miejsce, a Ares 216 i 267 miejsce na liście TOP500.
- Uruchomiono nową wersję strony internetowej ACK Cyfronet AGH.
- Rozpoczęto realizację projektów EuroHPC PL, PIONIER-LAB, KMD3, AGH – PANDA3, EGI ACE, EOSC Future i FINDR





- 2022** Uruchomiono superkomputer Athena o mocy obliczeniowej 7,7 PFlops.
Uroczyste otwarcie Data Center Podole.
Po raz pierwszy w historii polskiej informatyki trzy superkomputery z jednego polskiego centrum obliczeniowego (Athena, Ares i Prometheus) na liście TOP500.
Prometheus (53 748 rdzeni, 2,7 PFlops) po raz 15-ty na liście TOP500: 475 miejsce.
Ares (37 824 rdzenie, 4 PFlops) po raz 3-ci i 4-ty na liście TOP500: 290 i 323 miejsce.
Athena (6 144 rdzeni, 7,7 PFlops) po raz 1-szy i 2-gi na liście TOP500: 105 i 113 miejsce.
Inauguracja superkomputera LUMI.
Rozpoczęto realizację projektów: EUMaster4HPC, FAIRCORE4EOSC, EuroScienceGateway, InterTwin, DT-GEO, Geo-INQUIRE i EDITH
- 2023** Uroczyste obchody 50-lecia Cyfronet AGH z udziałem Prezydenta Andrzeja Dudy, który wręczył odznaczenia państwowe zasłużonym pracownikom Cyfronetu.
Athena zajmuje 123 miejsce, a Ares 362 miejsce na liście TOP500 (edycja czerwcową).
Cyfronet AGH uhonorowany nagrodą Polonia Minor.
Zespół projektu EuroHPC PL laureatem Polskiej Nagrody Inteligentnego Rozwoju 2023.
Po raz drugi w historii polskiej informatyki trzy superkomputery z jednego polskiego centrum obliczeniowego (Athena, Helios CPU i Ares) na liście TOP500 (edycja listopadowa), odpowiednio na miejscach: 154, 290 i 403.
Rozpoczęto realizację projektów: DOME, EUROCC 2, EUreka3D, PIONIER-Q i GEMINI
- 2024** Po raz pierwszy w historii polskiej informatyki cztery superkomputery z jednego polskiego centrum obliczeniowego (Helios GPU, Athena, Helios CPU i Ares) na liście TOP500 (edycja czerwcową), odpowiednio na miejscach: 55, 177, 305 i 442.
Helios GPU na 3. miejscu listy Green500 (edycja czerwcową).
Cztery superkomputery z Cyfronetu (Helios GPU, Athena, Helios CPU i Ares) ponownie na liście TOP500 (edycja listopadowa), odpowiednio na miejscach: 69, 212, 348 i 490.
Rozpoczęto realizację projektów: Meetween, EOSC Beyond, EPICURE, FFplus, PLGrid ICON i EPOS ON

KAMPUS AGH

LEGENDA

1. Rektorat
2. Wydział Inżynierii Ładowej i Gospodarki Zasobami
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. Studium Języków Obcych
22. Studium Wychowania Fizycznego i Sportu
23. Basen AGH
24. Centrum e-Learningu
25. Akademickie Centrum Komputerowe CYFRONET AGH
26. Centrum Rozwiązań Informatycznych
27. Centrum Organizacji Kształcenia
28. Centrum Spraw Studenckich
29. Szkoła Doktorska AGH
30. Dział Studentów Zagranicznych
31. Centrum Rekrutacji
32. Centrum Międzynarodowej Promocji Technologii i Edukacji AGH – UNESCO
33. Miasteczko Studenckie AGH
34. Uczelniana Rada Samorządu Studentów
35. Centrum Karier
36. Centrum Transferu Technologii
37. Dział Współpracy z Administracją i Gospodarką
38. Krakowskie Centrum Innowacyjnych Technologii INNOAGH
39. Centrum Obsługi Projektów
40. Centrum Obsługi Nauki
41. Centrum Technologii Kosmicznych
42. Centrum Spraw Międzynarodowych
43. Biuro ds. Osób Niepełnosprawnych
44. Muzeum Geologiczne WGGIOŚ
45. Wydawnictwa AGH
46. Akademickie Centrum Kultury Klub STUDIO
47. Klub Studencki Gwarek
48. Klub Studencki Zaścianek
49. Klub Studencki Filutek
50. Studio Muzyczne Kotłownia



ISBN 978-83-61433-43-9



Akademickie Centrum Komputerowe
CYFRONET AGH
ul. Nawojki 11

30-950 Kraków, P.O.Box 6
telefon: +48 12 632 33 55
fax: +48 12 633 80 54

cyfronet@cyfronet.pl
www.cyfronet.pl