Academic Computer Centre

CYFRONET AGH

2024

ACC CYFRONET AGH is a leading unit empowered by the Committee for Scientific Research to develop and manage the High-Performance Computers (HPCs) and Cracow Metropolitan Area Network (MAN). CYFRONET is the coordinator of the PLGrid Program and is recognized by the National Centre for Research and Development as a Centre of Excellence in the area of grid and cloud services.

Dear Readers!

Cyfronet's 50th anniversary has recently passed. It was an opportunity to reflect on the Centre's current and future activities. At the same time, it was a period of extremely intense work to provide our Users with the latest technologies supporting the development of Polish science, innovative economy and information society.

During this anniversary, we installed Helios in Cyfronet – a new supercomputer with a computing power five times greater than the power of our Athena, the fastest supercomputer in Poland. We are convinced that having approximately 35 PFlops, Helios will be among the European leaders in computing power.

The Helios supercomputer is the most important, although not the only, achievement of the EuroHPC PL project – National Supercomputing Infrastructure for EuroHPC. The project also provided access to world-class quantum annealing computing resources as part of a hybrid computing platform combining classical supercomputing (HPC) with quantum computing (QC). Other specialised platforms have also been implemented with respect to the needs of personalised medicine, radiotherapy, or energy efficiency of calculations. In addition, many IT tools have been created for use in various fields of science and economy.

The EuroHPC PL project, coordinated by Cyfronet, is a successful closing of the 50th anniversary. However, simultaneously, we enter the next half-century of activity by actively participating in many other initiatives. Cyfronet is organising further competitions for annual access to calculations on the LUMI supercomputer. In 2023, LUMI was the third supercomputer in the world in terms of computing power and also the third in terms of energy efficiency. Moreover, thanks to Poland's representation in the international LUMI-Q consortium, the Helios supercomputer will be directly connected to the quantum supercomputer built in the Czech Republic. We also must remember the continuous operation of the existing supercomputers: Prometheus, Ares and Athena, still the fastest supercomputer in Poland.

All mentioned resources are available within the PLGrid infrastructure initiated and coordinated by Cyfronet. The strategy of unification of access to various resources within the PLGrid infrastructure, undertaken years ago, has brought many benefits to users. After logging in to the PLGrid Portal, one can easily choose from many machines, virtual environments and software packages. The Helpdesk team, i.e. the Operations Centre, provides precious support. Cyfronet also organises training for users on effectively using the tools provided. In this regard, it is worth mentioning that Cyfronet coordinates the National HPC Competence Centre, which serves as a contact point for representatives of the innovative economy.

We also care about supporting the development of Open Science by actively participating in a series of EOSC projects and developing the so-called Digital Twins in medicine and Earth sciences. You can get more information about these and many other activities of Cyfronet on the following pages of this publication, which I wholeheartedly invite you to read. At the same time, I would like to thank all Cyfronet Friends and Users for their cooperation and valuable hints on the further development of the Centre. I constantly invite you to personal contact with Cyfronet employees.

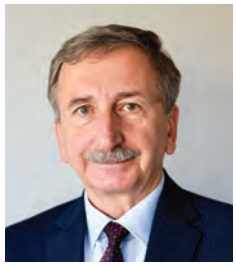
Yours sincerely,

Prof. Kazimierz Wiatr
Director of ACC Cyfronet AGH

DIRECTOR
Kazimierz Wiatr, Prof.
phone: +48 12 633 34 26
e-mail: Kazimierz.Wiatr@cyfronet.pl

Deputy Director for IT Infrastructure
Karol Krawentek, MSc Eng.
phone: +48 12 633 34 26
e-mail: Karol.Krawentek@cyfronet.pl

Deputy Director for High Performance Computers
Marek Magryś, MSc
phone: +48 12 633 34 26
e-mail: Marek.Magrys@cyfronet.pl



Deputy Director for Security and Innovative Economy
Henryk Baniowski, MSc Eng.
phone. +48 12 633 34 26
e-mail: Henryk.Baniowski@cyfronet.pl

Deputy Director for Administrative Affairs
Agnieszka Szymańska, MSc
phone: +48 12 633 34 26
e-mail: Agnieszka.Szymanska@cyfronet.pl

Deputy Director for Financial Affairs
Angelika Zaleska-Walterbach, MSc
phone: +48 12 633 80 53
e-mail: Angelika.Zaleska@cyfronet.pl

Computer Networks Department – Krzysztof Gaweł, MSc (manager)
Phone: +48 12 634 10 25, e-mail: Krzysztof.Gawel@cyfronet.pl

High Performance Computers Department – Patryk Lasoń, MSc Eng. (manager)
Phone: +48 12 632 33 55, e-mail: Patryk.Lason@cyfronet.pl

Data Security Department – Grzegorz Sułkowski, PhD Eng. (manager)
Phone: +48 12 632 33 55, e-mail: Grzegorz.Sulkowski@cyfronet.pl

Data Storage Department – Adrian Marszalik, MSc Eng. (manager)
Phone: +48 12 632 33 55, e-mail: Adrian.Marszalik@cyfronet.pl

HPC Software Department – Łukasz Flis, MSc (manager)
Phone: +48 12 632 33 55, e-mail: Lukasz.Flis@cyfronet.pl

Cybersecurity Department – Krzysztof Niziołek, MSc Eng. (manager)
Phone: +48 12 632 33 55, e-mail: Krzysztof.Niziolek@cyfronet.pl

Operational Center Department – Andrzej Zemła, PhD (manager)
Phone +48 12 632 33 55, e-mail: Andrzej.Zemla@cyfronet.pl

Users Department – Robert Pająk, MSc Eng. (manager)
Phone: +48 12 632 33 55, e-mail: Robert.Pajak@cyfronet.pl

Technical Department Nawojki – Damian Trela, MSc Eng. (manager)
Phone: +48 12 632 33 55, e-mail: Damian.Trela@cyfronet.pl

Technical Department Podole – Mariusz Kula, MSc Eng. (manager)
Phone: +48 12 632 33 55, e-mail: Mariusz.Kula@cyfronet.pl

Administrative Department – Agnieszka Szymańska, MSc (manager)
Phone: +48 12 633 34 26, e-mail: Agnieszka.Szymanska@cyfronet.pl

Financial Department – Angelika Zaleska-Walterbach, MSc (manager)
Phone: +48 12 633 80 53, e-mail: Angelika.Zaleska@cyfronet.pl

Projects Department – Aleksandra Pałuk, MSc (manager)
Phone: +48 12 632 33 55, e-mail: Aleksandra.Paluk@cyfronet.pl

Laboratory of the PLGrid Program – Jacek Kitowski, Prof. (manager)
Phone: +48 12 633 34 26, e-mail: kito@agh.edu.pl

Laboratory of Computing Acceleration and Artificial Intelligence
Paweł Russek, PhD DSc Eng., Associate Professor (manager)
Phone: +48 12 633 34 26, e-mail: Pawel.Russek@cyfronet.pl

Laboratory of Cloud Technologies – Łukasz Dutka, PhD (manager)
Phone: +48 12 633 34 26, e-mail: Lukasz.Dutka@cyfronet.pl

Laboratory of Data Processing – Roksana Wilk, Eng. (manager)
Phone: +48 12 632 33 55, e-mail: Roksana.Wilk@cyfronet.pl

Laboratory of Interdisciplinary Scientific Computing – Joanna Kocot, MSc Eng. (manager)
Phone: +48 12 632 33 55, e-mail: Joanna.Kocot@cyfronet.pl

Laboratory of Quantum Computing – Mariusz Sterzel, PhD (manager)
Phone: +48 12 632 33 55, e-mail: Mariusz.Sterzel@cyfronet.pl

Laboratory of Informatics Methods in Medicine – Marian Bubak, PhD Eng. (manager)
Phone: +48 12 633 34 26, e-mail: bubak@agh.edu.pl

Laboratory of Parallel Algorithms – Bogusław Cyganek, Prof. (manager)
Phone: +48 12 632 33 55, e-mail: cyganek@agh.edu.pl

Laboratory of Applications of Computational Techniques
Łukasz Rauch, PhD DSc Eng., Associate Professor (manager)
Phone: +48 12 632 33 55, e-mail: lrauch@agh.edu.pl

Laboratory of Visual Techniques – Jacek Przybylski, MSc (manager)
Phone: +48 12 632 33 55, e-mail: Jacek.Przybylski@cyfronet.pl


CYFRONET AGH

Term of office: 2021-2024

Marek Jarnicki, Prof. *The Chairman*	Jagiellonian University
Tadeusz Lesiak, Prof. *Deputy Chairman*	Institute of Nuclear Physics Polish Academy of Sciences
Jakub Badowski, MSc Eng.	Oil and Gas Institute National Research Institute
Marek Gorgoń, Prof.	AGH University
Jarosław Górski, MSc Eng.	Academy of Physical Education
Robert Gryboś, PhD	Jerzy Haber Institute of Catalysis and Surface Chemistry Polish Academy of Sciences
Tomasz Hachaj, Prof.	University of the National Education Commission
Marek Kisiel-Dorohinicki, Prof.	AGH University
Mariusz Kwinta-Pudełko, Msc Eng.	University of Agriculture
Wojciech Majka, MSc Eng.	Jagiellonian University Medical College
Tomasz Masier, MSc Eng.	Institute of Forensic Expertise
Roman Mazur SDB, PhD	The Pontifical University of John Paul II
Paweł Mikłaszewski, MSc	Academy of Music
Wojciech Palacz, PhD	Jagiellonian University
Dariusz Put, Prof.	Krakow University of Economics
Elżbieta Tomczyk, MSc	Łukasiewicz Research Network Krakow Institute of Technology
Dariusz Żelasko, PhD Eng.	Cracow University of Technology
Jacek Barski, Deputy Director	Marshal Office of the Małopolska District
Paweł Schmidt, Director	Cracow City Hall

Athena – strong support for scientific calculations

Athena achieves the theoretical computing power of over **7.7 PFlops**, which gave the machine 123rd place on the TOP500 list in June 2023 and made it the fastest supercomputer in Poland at that time. The new, accelerated computing system, installed in Cyfronet in 2021, provides the Polish scientific community and innovative economy with the most modern computing resources based on the latest generation GPGPU processors and accelerators along with the necessary data storage subsystem based on very fast flash memories.

Athena in numbers	
Number of computing cores	6 144
Number of GPGPUs	384
Computing power	7.7 PFlops
TOP500 – the list of the world's fastest computers (June 2023 edition)	123rd position

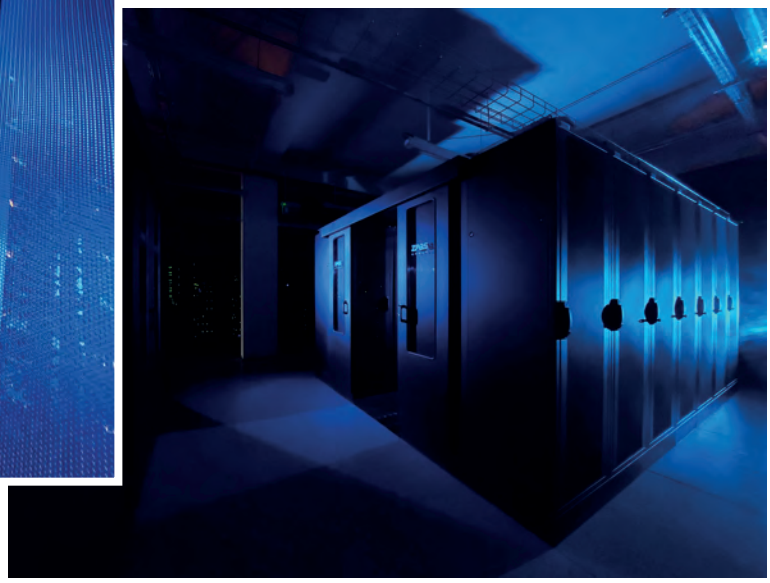
Athena's configuration includes 48 servers with AMD EPYC processors and 1 TB of RAM (6,144 CPU compute cores in total) as well as 384 NVIDIA A100 GPGPU cards.

The indispensable element enabling the use of such high computing power in an effective way is the provision of a high-performance internal network of a supercomputer (Infiniband HDR with 4 x 200 Gb/s bandwidth per server) and a very fast disk subsystem. It is built on the basis of the Lustre open source software, currently used in Ares and Prometheus supercomputers, and dedicated disk servers equipped with flash memory in the NVMe standard. The system was installed in the existing Cyfronet data centre and integrated with the PLGrid infrastructure.

This type of infrastructure meets the needs of users of Cyfronet supercomputers, who use the computing infrastructure both to perform standard high-performance scientific simulations (HPC) and to apply artificial intelligence (AI) and machine learning (ML) methods to conduct research in the field of medicine, pharmacology, biology, chemistry, physics and many other fields of science. **Athena's computing power for AI computing is over 240 PFlops!**

The expected effect of delivering specialised computational resources of Athena will be the extension of the scope of research works, the possibility of undertaking advanced simulations and analyses, and increasing the possibilities of processing continuously flowing data from laboratories worldwide. The direct expected results of the work will be articles and scientific studies, patents, and in the longer term, the innovative solutions that may be the basis for the development of new solutions in the economy.


ATHENA
ATHENA
ZPAS
GROUP
AGH



Ares – towards shorter computation time

In 2021, the Ares supercomputer was launched in Cyfronet. It is built of computing servers with Intel Xeon Platinum processors, divided into three groups:

- 532 servers, each equipped with 192 GB of RAM,
- 256 servers, each equipped with 384 GB of RAM,
- 9 servers, each with 8 NVIDIA Tesla V100 cards.

Ares in numbers	
Number of computing cores	37 824
RAM	147.7 TB
Number of GPGPUs	72
Computing power	4 PFlops
TOP500 – the list of the world's fastest computers (June 2023 edition)	362nd position

The total theoretical performance of the CPU parts is over 3.5 PFlops, and the GPU part is over 500 TFlops. Ares is also supported by a disk system with a capacity of over 11 PB. An InfiniBand EDR network is used for data transfer. The supercomputer has 37,824 computing cores and 147.7 TB of RAM. It is also equipped with a liquid cooling system.

Ares complements Cyfronet's computing resources by providing a newer generation of processors and servers with more memory. It enables shortening the computation time of scientific tasks and addressing problems that so far could not be run on a large scale due to insufficient memory. In addition, placing Ares in Data Center Podole, a geographically different location than Prometheus, guarantees the continuity of the provision of computing services in crisis situations.

Prometheus – PetaFlops computing power

Changes in the world of science follow very quickly and affect the speed of development of IT facilities, which Cyfronet offers to scientists. Researchers' growing demands for computing power and data storage are clearly visible from the disciplines almost traditionally associated with high-performance computers: chemistry, physics, astronomy, life sciences and fields related to them. Astronomy, astrophysics and space physics are based on the one hand on data acquisition and analysis, and on the other on complex computer simulations. Biological, chemical and medical sciences as well as those mentioned above are characterized by rapid development and introduction of new, increasingly sophisticated research methods, e.g. molecular techniques based on high-performance DNA sequencing. Medicine, as a multidisciplinary field, deals with a number of time-consuming analyses, e.g., the human genome. It results in increased demand for automated collection, storage and analysis of biomedical signals and images, what in turn leads to necessity of use of the supercomputing resources in order to implement these processes. The possibility of linking together multiple unique data, i.e. the clinical, genetic as well as environmental and social data, brings many benefits, but also in this case the dedicated services are needed that can be offered only by supercomputing centers.

These are the tasks Prometheus supercomputer deals with. As the successor of Zeus, it has become a part of the PLGrid infrastructure and serves scientists, also within international research projects. Prometheus is used for: data results analysis, numerical simulations, (big) data processing, and advanced visualisations provision.

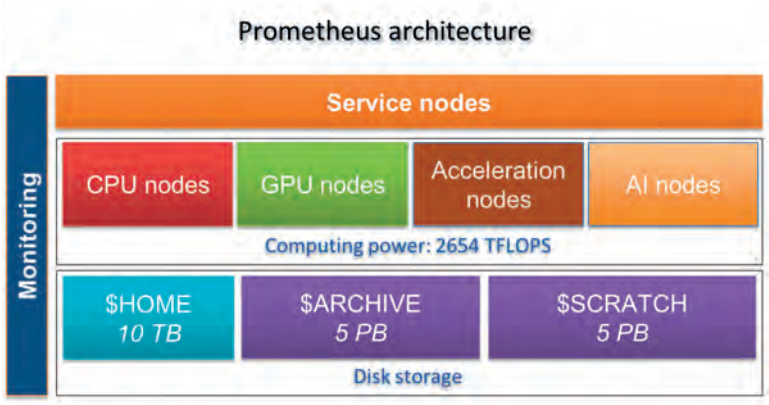
Prometheus consists of more than 2,239 servers based on the HP Apollo 8000 platform, combined with the super-fast InfiniBand FDR network with 56 Gbit/s capacity. Its energy saving and high-performance Intel Haswell and Intel Skylake processors offer 53,748 cores. These are accompanied by 283.5 TB of DDR4 RAM and by two storage file systems of 10 PB total capacity, and 180 GB/s access speed. Prometheus has also been equipped with 144 NVIDIA Tesla K40 XL and 32 NVIDIA Tesla V100 GPGPUs. The theoretical performance of Prometheus is 2.7 PFlops!

Due to the innovative technology of direct liquid cooling of processors and RAM modules, Prometheus is also one of the most energy-efficient computers in its class in the world. This was achieved by using the cooling water having a temperature of 28°C. To cool down the water to such a temperature in our climate it is enough to use cheap in use dry-coolers, instead

of ice water generators, consuming relatively large amounts of electricity. With use of water cooling, electronic components operate at temperatures lower than normal, what positively affects not only the failure, but also allows to reach efficiency more than 5% higher than for a similar installation based on the classic air cooling. Furthermore, liquid cooling allowed for extremely high installation density of 144 computing servers in one rack, therefore Prometheus, weighing of more than 40 tons, covers 18 m^2 area and is placed on 20 racks only. This also has a significant impact on internal data transmission, because distances of connections are critical here.

Prometheus in numbers	
Number of computing cores	53 748
RAM	283.5 TB
Number of GPGPUs	144
Computing power	2.7 PFlops

Prometheus has been installed in a high-tech computing room, exclusively adapted for its operation. The supercomputer's proper functioning is additionally supported by the accompanying infrastructure, including such systems as guaranteed power supply with an additional generator, modern air-conditioning and gas extinguishing.



Since its installation in 2015, Prometheus has been continuously listed (15 times) on the TOP500 list, occupying high places, with **38 –** the highest noted spot. Prometheus was the fastest supercomputer in Poland 11 times.

The Prometheus architecture is a composite of several classes of nodes, varying in terms of architecture of computing resources and functionality:

- classical cluster of computing servers with highly efficient CPU nodes equipped with two Intel Xeon processors,
- cluster of servers equipped with graphic accelerators GPGPU NVIDIA Tesla K40 XL,
- acceleration partition with a set of devices supporting the Prometheus configuration with several types of accelerators (including GPGPU NVIDIA K80, Intel Xeon Phi 7120P, and Nallatech FPGA cards),
- a partition dedicated to calculations related to artificial intelligence, equipped with GPGPU NVIDIA Tesla V100 graphics accelerators. It is worth mentioning that **this partition is a system with computing power over 4 PFlops for tensor operations and 256 TFlops for standard calculations performed on double precision numbers, which makes it one of the fastest dedicated solutions for artificial intelligence available for the needs of science in Poland.**

Thanks to Prometheus users have received more than seven times greater opportunities compared to the previously used Zeus. Much more efficient processors, faster network of internal connections, and a greater amount of memory of Prometheus enable to perform calculations on a scale impossible to achieve using previous Cyfronet's resources.

Three supercomputers from Cyfronet together on the TOP500 list of the fastest supercomputers in the world

Zeus supercomputer

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

Zeus and Prometheus supercomputers

- 2015 - VII, 269th and 49th place, 374 + 1659 TFlops
- 2015 - XI, 387th and **38th** place, 374 + 2399 TFlops

Prometheus supercomputer

- 2016 - VI, 49th place, 2399 TFlops
- 2016 - XI, 60th place, 2399 TFlops
- 2017 - VI, 72nd place, 2399 TFlops
- 2017 - XI, 78th place, 2399 TFlops
- 2018 - VI, 103rd place, 2399 TFlops
- 2018 - XI, 131st place, 2399 TFlops
- 2019 - VI, 174th place, 2399 TFlops
- 2019 - XI, 241st place, 2399 TFlops
- 2020 - VI, 288th place, 2399 TFlops
- 2020 - XI, 324th place, 2399 TFlops

Prometheus and Ares supercomputers

- 2021 - VI, 373rd and **216th** place, 2399 + 3510 TFlops
- 2021 - XI, 440th and 267th place, 2399 + 3510 TFlops

Prometheus, Ares and Athena supercomputers

- 2022 - VI, 475th, 290th and **105th** place, 2399 + 3510 + 7709 TFlops

Ares and Athena supercomputers

- 2022 - XI, 323rd and 113th place, 3510 + 7709 TFlops
- 2023 - VI, 362nd and 123rd place, 3510 + 7709 TFlops

Supercomputers from Cyfronet on the TOP500 list


TOP500 CERTIFICATE

The List.

Athena - FormatServer THOR ERG21, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100 SXM4 40 GB, Infiniband HDR

Cyfronet, Poland

is ranked

No. 105

among the World's TOP500 Supercomputers

with 3.70 PFlop/s Linpack Performance

on the 59th TOP500 List published at the ISC22

Conference on June 01, 2022.

Congratulations from the TOP500 Editors

Erich Strohmaier NERSC/Berkeley Lab

Jack Dongarra University of Tennessee

Horst Simon NERSC/Berkeley Lab

Martin Meuer Prometeus

On May 30, 2022, during the ISC High-Performance conference in Hamburg, the next TOP500 list of supercomputers with the highest computing power in the world was announced. **For the first time in history, the list simultaneously included three supercomputers from one Polish computing centre. These were those operating in Cyfronet: Athena (105th place), Ares (290th), and Prometheus (475th).**

The story of the machines installed in Cyfronet that were on the TOP500 list began in 1996 when SPP1200/XA-32 computer took 408th place. After a long time without records, the Zeus supercomputer appeared on the list in 2010 and remained there until 2015. From then on, until 2022, Cyfronet has been represented by Prometheus, which was joined by Ares in June 2021 and by Athena in June 2022.

A list of places that Zeus, Prometheus, Ares and Athena supercomputers have taken on the TOP500 list in recent years is presented in the adjacent column.

The presence of Cyfronet supercomputers on the TOP500 list is an important confirmation of competences in the following areas: availability of resources, quality of services, users' trust, reliable operation and greater access to EU programs.

The world's top energy efficiency

In 2022, all three of Cyfronet's supercomputers that took places on the TOP500 list have also been ranked on the Green500 list of the most ecological supercomputers. The main criterion (energy efficiency) is calculated as the ratio of the number of floating-point operations per second (computing power of a supercomputer) to energy consumption: Gflops/W. Athena's **9th place** on the Green500 was a particular success. This position proves the excellent ratio of the provided computing power to the electricity consumption. In 2023, Athena and Ares were also listed on the Green500 list, taking 19th and 101st place respectively.


The GREEN 500 CERTIFICATE

Athena - FormatServer THOR ERG21, AMD EPYC 7742 64C 2.25GHz, NVIDIA A100 SXM4 40 GB, Infiniband HDR

Cyfronet, Poland

is ranked

No. 9

among the World's TOP500 Supercomputers

in the Green500 List published at the ISC22 Conference on June 01, 2022.

Congratulations from the Green500 Editors

Wu-chun Feng, Virginia Tech

Kirk Cameron, Virginia Tech

Supercomputers usage

Supercomputers in Cyfronet are part of the European cloud and grid infrastructure under the European Grid Infrastructure (EGI). At the same time, they are also important supercomputers in the PLGrid nationwide computing infrastructure – the platform for conducting *in silico* research and enabling calculations with use of high-performance computers, also within the cloud and grid architecture.

Via the PLGrid infrastructure scientists can get access to the supercomputers' resources. Dedicated computing environments, so-called domain grids, and specialised IT platforms enable conduction of increasingly complex research problems. The research portfolio carried out with the help of Cyfronet supercomputers is quite reach. It includes:

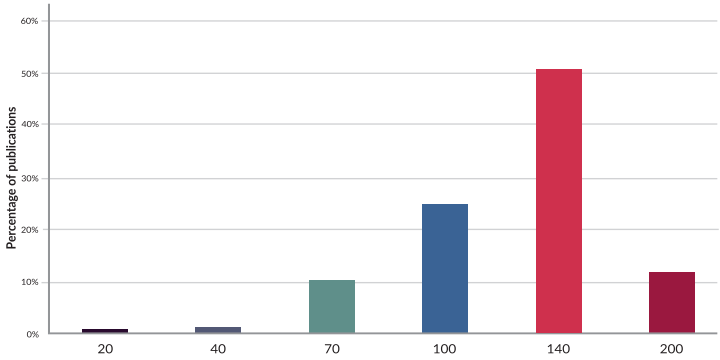
- production of models and tools enabling the construction of Digital Twins,
- measurement of the associated production cross-section of a Z boson with two beauty quarks in the unique kinematic region available at the LHCb experiment,
- recognition and diarisation of speakers using artificial neural networks,
- multi-agent systems for analysis and prediction of pandemic development in Poland,
- the influence of pharmacokinetic profiles on the effectiveness of depression treatment,
- Machine Learning techniques for the analysis of large amounts of data from astroplasma simulations,
- the analysis of gene expression in rainbow trout in response to Aeromonas infection,
- the use of polymers based on β-cyclodextrin derivatives in the elimination of water pollution with drugs,
- archaeogenetics of prehistoric human populations,
- examination of the structure of human kininogen by cryogenic electron microscopy,
- numerical calculations of air and methane flow in the longwall walkway and its surroundings,
- modeling the effectiveness of selected strategies for the protection of beneficial insects in agricultural landscapes of the EU.

A wide range of research topics is evidence of constantly increasing number of scientists, who are aware of advantages of supercomputers. With their help one can get the final results of huge simulations many, many times faster, compared to the case of an ordinary, desktop computer. Supercomputers enable to significantly reduce time of computations that using a single computer would often take many years (in specific cases more than 150, 700 or even 1000 years). Here they may be

usually performed within a few days. What is important, Cyfronet users can benefit from the professional support – starting from full documentation, through training, to individual consultations with experts.

Year	No. of jobs	CPU time in years
Zeus supercomputer		
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
Zeus and Prometheus altogether		
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
Zeus, Prometheus and Ares altogether		
2021	5 549 582	43 409
Zeus, Prometheus, Ares and Athena altogether		
2022	6 227 244	48 716

Ministry of Education and Science marks of articles published in 2022 by Cyfronet Users in scientific journals



In addition to individual scientists and small research groups, even international consortia carry out calculations from many different scientific disciplines with the help of supercomputers – of course with the participation of Polish scientists. Scientific computations do not include simulations only. Computing power is utilised by Polish researchers also within international scientific projects, including experiments like CTA, LOFAR, EPOS, Large Hadron Collider in CERN and the recently discovered gravitational waves in LIGO and VIRGO detectors.

Obviously, even the highest positions in the TOP500 list, or the latest technologies used to build high-performance computers do not fully reflect the importance of this kind of computing resources for the Polish scientific community. The usefulness of supercomputers provided by ACC Cyfronet AGH as a tool for conducting research is best evidenced by statistical data on their use.

The table presents the aggregated key data on the number of computational tasks and their duration, performed by Cyfronet for other units.

It is worth mentioning that huge users' demands for computing power and space for data storage would not be fulfilled without continuous extension of computing resources and disk storage. Therefore, we carefully analyse users' suggestions and statistical data related to carried out computations together with world's trends in computing.

The scientific level of the tasks carried out with the use of the infrastructure provided by ACC Cyfronet AGH is very high. This is evidenced by the results of scientific and research works carried out in 2022 using this infrastructure, which were presented in many publications.

Comprehensive infrastructure of efficient and safe storage of digital data

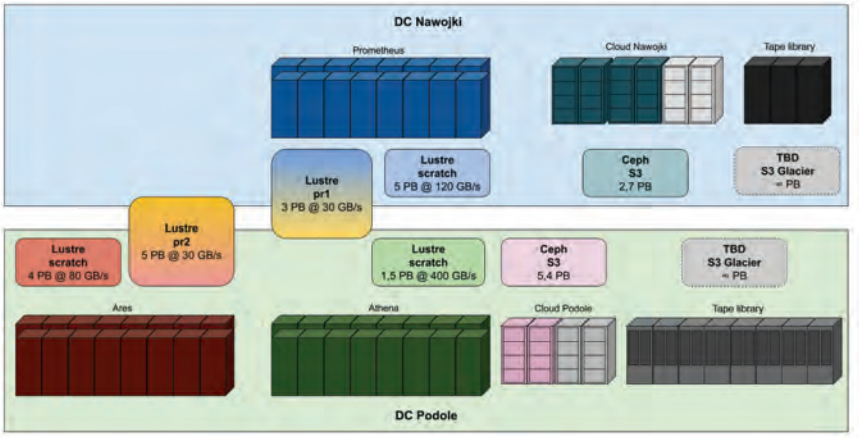
The currently observed phenomenon of the rapidly growing amount of digital information also applies to the scientific community. Access to very efficient supercomputers enables performing analyses of large-scale research problems, what results in generation of huge data sets. They require a completely new approach to information processing and storage. This problem, being currently one of the most important challenges of the modern digital world, is described by the concept of BigData. Also in ACC Cyfronet AGH there are clearly visible the growing expectations concerning available capacity, speed and additional functionalities of storage resources, caused by offering more efficient computing systems. The architecture of the Cyfronet Data Storage System, the main mass storage platform for High-Performance Computers, is composed of following elements:

- the SAN network – the efficient and highly available network dedicated to communication among devices within the Data Storage System, and clients using shared resources or services,
- disk arrays and servers of various types, offering the storage space for the users' data – starting from fast, but expensive and less capacious solutions, and ending with the devices with large storage capacity and relatively cheap, but with limited efficiency,
- service servers, with specialised tools and virtualisation software, providing users with functionalities such as automatic backup and archival, hierarchical data storage systems, high-performance hardware file platforms or distributed network file systems,
- tape libraries and specialised software used to store critical user data on magnetic media,
- additional infrastructure, including Ethernet, Infiniband as well as solutions supporting management of the IT infrastructure and enabling secure storage of magnetic media.

At present, the total storage capacity of Cyfronet disk and tape resources is approximately **75 PB.**

Mass storage for supercomputers

The proper teaming of computing infrastructure with the right selection of storage solutions can assure the best quality of services provided to scientific users. The scale of problems in this area grows with the complexity and efficiency of the supercomputers used. Currently, data storage systems attached to Cyfronet supercomputers store billions of files up to terabytes. The broad thematic scope of research on the resources provided by Cyfronet is reflected in the variety of configurations of the Centre's key supercomputers and thus also in the structure of dedicated storage resources. The resources of the Cyfronet data storage system are located in two locations.



Supercomputers use, among others, efficient temporary space, the so-called scratch. The critical element here is the speed of operation, which is why it is based on a high-speed distributed file system architecture – Lustre. The advantage of Lustre is the ability to scale the capacity and efficiency of the disk space. By combining the capacity of multiple servers, I/O bandwidth is aggregated and scales with additional servers. Moreover, bandwidth and/or capacity can be easily increased by dynamically adding more servers without interrupting users' computations. Currently, all supercomputers in Cyfronet use the scratch space implemented by Lustre. In the case of Prometheus, this space has the capacity of 5 PB and the speed of 120 GB/s. Ares has the space with the total of 4 PB and the speed of 80 GB/s. In both of these computers, the scratch space is realized with the help of mechanical disks. In the case of Athena, user data is stored on solid-state drives. The use of this type of solution significantly increases the efficiency of the system. The planned capacity of this type of space for Athena is to be 1.5 PB and achieve the bandwidth of 400 GB/s.

Most of Cyfronet's disk memory resources are dedicated to the needs of users of domain services developed in the PLGrid program. The PLGrid infrastructure offers a dedicated workspace for groups using domain services – the functionality necessary to enable collaboration between scientists working in geographically dispersed locations. This functionality is implemented using the Lustre file system. The maximum capacity of the /pr1 resource in the Prometheus supercomputer is 5 PB, and the total speed of reading and writing operations reaches 30 GB/s. In the case of Ares and Athena, the /pr2 resource has the capacity of 5 PB and the speed of 30 GB/s.

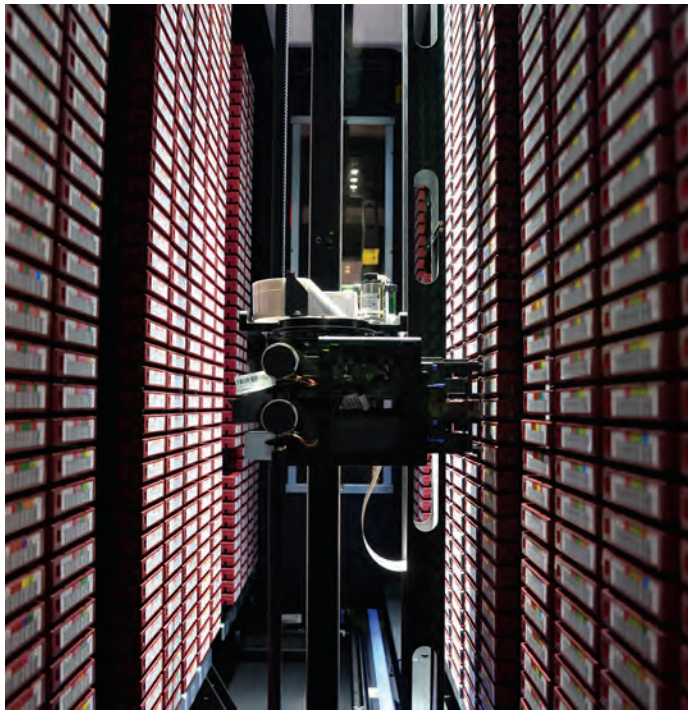
The object-oriented data storage system is an additional resource for storing users and projects' data resources in Cyfronet. It is based on the CEPH software. The data in this system is available through the S3 protocol based on the REST API and is stored in globally unique containers (buckets) in which users store their data in the form of objects.

A particular case of mass storage are resources for large projects and international collaborations in which Cyfronet participates, such as WLCG (Worldwide LHC Computing Grid), analyzing data from the LHC detector at CERN, or CTA (Cherenkov Telescope Array), studying gamma radiation using a network of radio telescopes. These projects require substantial disk resources, often available using unusual protocols such as SRM, xroot, or GridFTP. Cyfronet provides this type of disk space using several instances of dedicated DPM software (Disk Pool Manager) and using dedicated networks such as LHCone. The total capacity of DPM systems in Cyfronet exceeds 2 PB.

Currently, the total available disk capacity used by ACC Cyfronet AGH is approximately **45 PB.**

Backup-archiving services in detail

ACC Cyfronet AGH provides its users with a wide portfolio of services related to securing information stored in a digital form. In addition to advanced technological solutions such as communication networks dedicated to storage systems, modern disk arrays or hardware file servers, the Centre also performs conventional backup-archiving services, based on magnetic media. Contrary to the expectations of the inevitable end of solutions using data storage on magnetic tapes, this technology is constantly evolving, and offers in successive generations not only the increasing capacity of the media, but also significantly better capacities and mechanisms supporting the safety and effectiveness of the information storage (e.g. data encrypting and compressing algorithms, which are embedded in the tape drives).



Cyfronet has currently three tape libraries having in total over 14 thousand slots for LTO magnetic tape drives and 34 drives of the VI, VII and IX generation. A single LTO-7 magnetic medium has a physical capacity of 6 TB and allows recording at the speed up to 300 MB/s, which theoretically allows the storage of almost 190 PB of uncompressed data in tape libraries. Described resources are used for performing current backup and archive of important information resources of the Centre's users.

Backup is performed on the active data – that might be currently in use – through a replication process from the source location to a separate, isolated destination. The ideal backup procedure ensures consistency of the source and backup data, both at the level of a single object (a file located on a hard drive), and in the case of complex IT systems, such as database or mail servers as well as virtual environments. Physically, the cloning process is usually done by copying the source data from the backup client disk to disk/tape resources of the target backup server, using a dedicated or shared access medium, such as Ethernet or SAN. The purpose of an archive is to ensure security of unused data and to release occupied storage resources. In contrast to the backup, the archive is performed once, by the migration of the data from the source location to the destination.

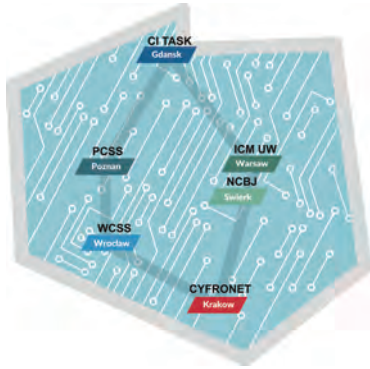
ACC Cyfronet AGH provides a wide range of backup services, addressed directly to users, and operating without their interaction. Among those at the disposal of users, there are ones based on FTP, NFS and SCP network protocols, acting within the dedicated backup servers. These machines provide backup solutions for users, allowing them to direct access to the backup data. It is up to users to decide which data they treat as a backup and which as archives.

At present, the total storage capacity of Cyfronet tape resources is approximately **30 PB**.

PLGrid Program

The infrastructure and the projects

The idea of the PLGrid Program has been invented by Cyfronet as a result of vast knowledge and experience gained in variety of national and EU projects. In 2007, it constituted formally as the PLGrid Consortium. At present, the Consortium consists of the following institutions: Academic Computer Centre in Gdansk, Interdisciplinary Centre for Mathematical and Computational Modelling in Warsaw, National Centre for Nuclear Research in Otwock-Świerk, Poznan Supercomputing and Networking Center, Wroclaw Centre for Networking and Supercomputing, and **Academic Computer Centre Cyfronet AGH as the initiator and leader of the PLGrid Program and Consortium**. The work carried out by Consortium partners led to the full-fledged distributed infrastructure for scientific computing. This infrastructure comprises not only high performance computing hardware, but also mass storage and dedicated tools for deployment of scientific applications on the available resources.



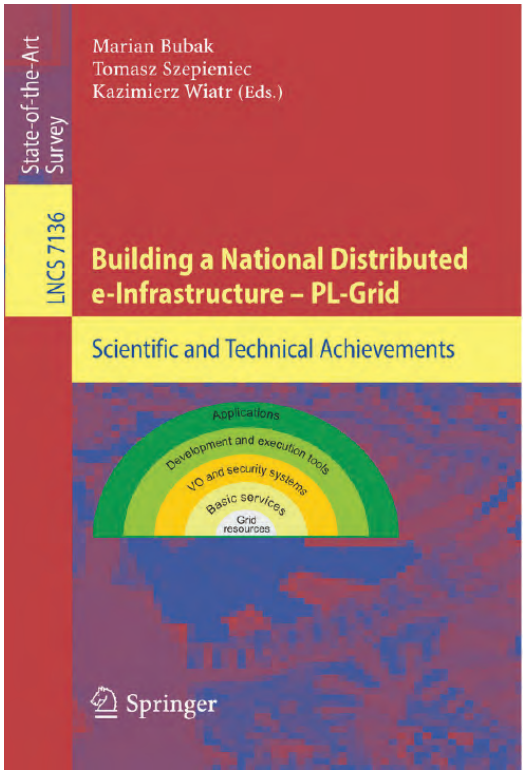
The design and construction of the PLGrid infrastructure started in the framework of the PL-Grid project (Polish Infrastructure for Supporting Computational Science in the European Research Space), in response to science needs, in which computers become more and more important. The main goal of the built infrastructure was to support scientists' investigations by integrating experimental data and results of advanced computer simulations carried out by geographically distributed research teams with use of supercomputers localised in High Performance Computing Centres. This aim was accomplished, among others, by extending the amount of computational resources in all PLGrid Consortium institutions. What is more, thanks to the PL-Grid project, in fall 2011 all Consortium partners have been spotted on TOP500 – the list of fastest world supercomputers. The same year Zeus supercomputer in Cyfronet has been located at 80th position – what gave it the first place among Polish supercomputers.

The next step of the PLGrid Program was to provide the researchers with necessary IT support through preparation of the specific computing environments, i.e., services and software as well as helping users in planning, running and analysing complex scientific experiments. Preparation of dedicated computing environments, so called domain grids, tailored to the needs of 13 different groups of scientists, was the most important task of PL-Grid follow-up – implemented within the PLGrid Plus project (Domain-oriented services and resources of Polish Infrastructure for Supporting Computational Science in European Research Space).

Adaptation of the infrastructure to the needs of scientists brought by domain grids was a great success of the PLGrid Plus project. Therefore, these activities have been further extended by the "New generation domain-specific services in the PL-Grid Infrastructure for Polish Science" project. In the PLGrid NG project, the domain-specific grids were developed for several other groups of scientists, representing fourteen research fields **(in total, in the two projects, IT support tools were built for 27 scientific disciplines)**.

However, the PLGrid Program did not stop on development of domain-oriented solutions only. Thanks to longstanding involvement in the development of grid computing infrastructures, Cyfronet is now recognised as a Centre of Excellence in the area of cloud and grid services – an achievement reflected by the new large-scale scientific grant named Distributed Computer and Data Infrastructure Centre of Excellence – PLGrid Core. This grant represented the next step in the development of the PLGrid Program and extension of the infrastructure towards Cloud Computing and handling big data calculations. It aimed not only at extension of hardware and software portfolio, but also dedicated accompanying facilities. One of them – a new backup Data Center built in separate geographical location highly increased security of scientific data sets.

It is worth noting that on the November 2015 edition of TOP500 the **Prometheus supercomputer, deployed at Cyfronet in 2015 in the framework of PLGrid Core, took the 38th position, the highest so far for supercomputers deployed in Poland!**

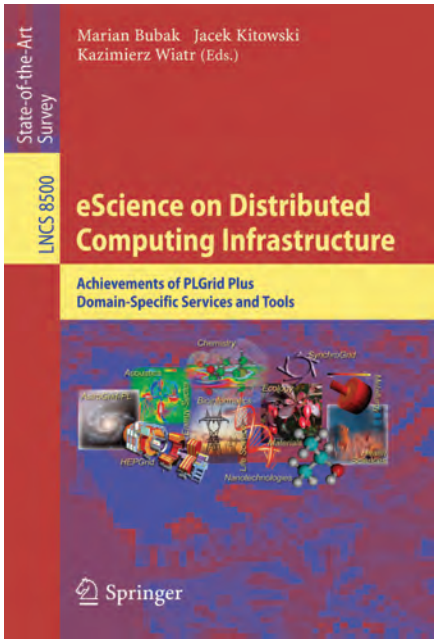


At present, enormous computing power and mass storage resources are available within the infrastructure. In addition, many tools supporting organization of computational experiments, designing and running applications, computationally supporting research and results' visualisation were implemented in the infrastructure. Furthermore, the Consortium offers Cloud Computing.

All the projects of the PLGrid Program have been co-funded by the European Regional Development Fund as part of the Innovative Economy program. ACC Cyfronet AGH has the honour to be their responsible coordinator.

Domain-specific grids in the PLGrid infrastructure

The PLGrid infrastructure, established through the Cyfronet's initiative, offers a uniform access to resources of all Polish High-Performance Computing centres. Unification takes place at many levels, ranging from a user's single login and password across the infrastructure, to the access to scientific applications. Sometimes, however, the use of modern computing systems, services and tools of the e-infrastructure becomes relatively difficult for researchers. Basic infrastructure services are often insufficient to conduct scientific research, particularly in the context of large international consortia. In such situations, users need both assistance and close collaboration with service providers.



Therefore, within the PLGrid Plus project (2011-2015), the PLGrid infrastructure has been extended with specific environments, solutions and services, developed according to the identified needs of 13 pilot groups of scientists. The main aim of the project was to lower the barriers required for researchers to use the infrastructure, and, thus, attract new communities of users, who need the computational power and large disk space of supercomputers, but have no or little skills in using it. To enable and facilitate development of domain--specific environments, the project relied on a broad cooperation with representatives of various disciplines, often grouped in domain consortia.

The dedicated services hide the complexity of the underlying infrastructure and, at the same time, expose the actual functions that are important to researchers of the given domain. In this way, users are provided with exactly the functionality they need. What is more, it is exposed to them in their domain-specific manner to achieve maximum intuitiveness and usefulness.

Scientific and technical achievements of PLGrid Plus were presented in a book published in the Springer Publisher, in September 2014. The book is an important source of information for researchers, developers and system administrators, who use grid and cloud environments in their research. The book contains 36 chapters and is divided into three parts: the first one (chapters 1 to 8) provides a general overview of the work carried out in the project and a description of the current state of the PLGrid infrastructure, including new solutions in the field of security and middleware.

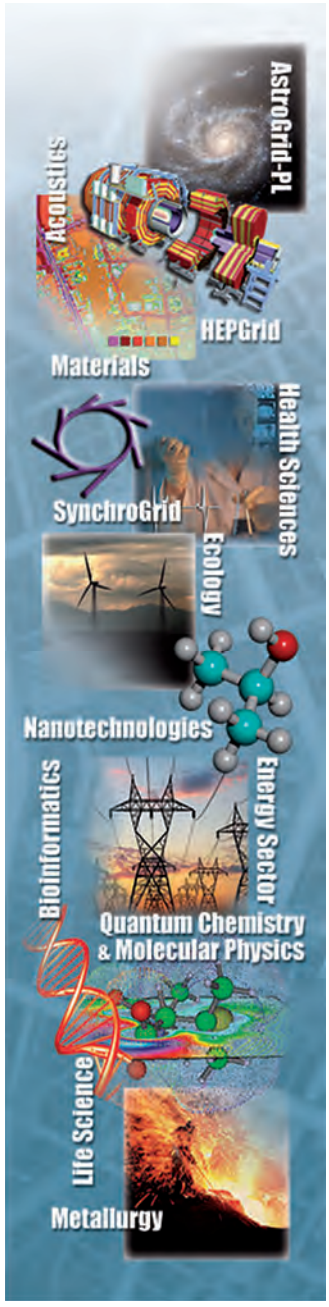
The second part (chapters 9 to 13) presents new environments and IT services that can be used by all of the previously mentioned groups of scientists. The third part (chapters 14 to 36) describes how specific environments, tools and services, prepared within the PLGrid Plus project, are used in advanced computations and computer simulations performed by different groups of researchers. These chapters present computational models, new algorithms and methods of their implementation using available tools and services.

Success of the PLGrid Plus project, in particular, the growing popularity of specialised tools and platforms prepared for the members of the first 13 strategic areas of science, led to a rapid increase in demand for related services to researchers in other fields. Therefore, the PLGrid Consortium launched the PLGrid NG project (2014-2015), whose primary objective was to implement, within the PLGrid infrastructure, several additional computing services for groups of scientists representing 14 new research fields.

New domain-specific services covered a wide range of activities: including provision of the specialised software, mechanisms of data storage and modern platforms integrated with a new type of tools and dedicated databases, which sped up research conduction as well as streamlined and automated the work of research groups.

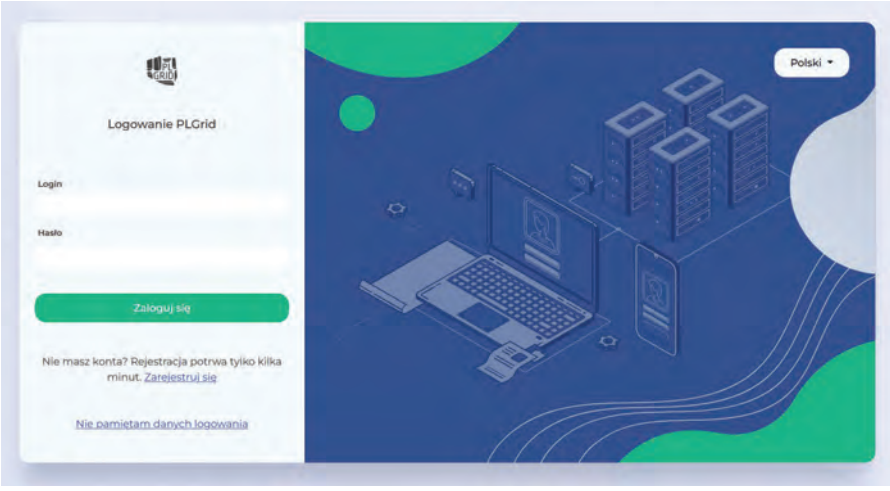
Preparation and implementation of a set of domain-specific services fit very well with the need of development of an advanced IT infrastructure designed for the implementation of modern scientific research. The well-tailored PLGrid e-infrastructure does not only fulfil researchers' needs for suitable computational resources and services, but also enables Polish scientific units collaboration with international research organizations.

Expansion of the existing computational infrastructure towards domain-specific solutions for research teams allowed more effective research conduction.



Access to resources

Access to the PLGrid infrastructure resources is obtained by registering in the **PLGrid Portal** (*https://portal.plgrid.pl/*). Access to huge computational power, large storage resources and sophisticated services on a global level is free to Polish researchers and all those engaged in scientific activities associated with the university or research institute in Poland. All one has to do is to create an account via the PLGrid Portal.



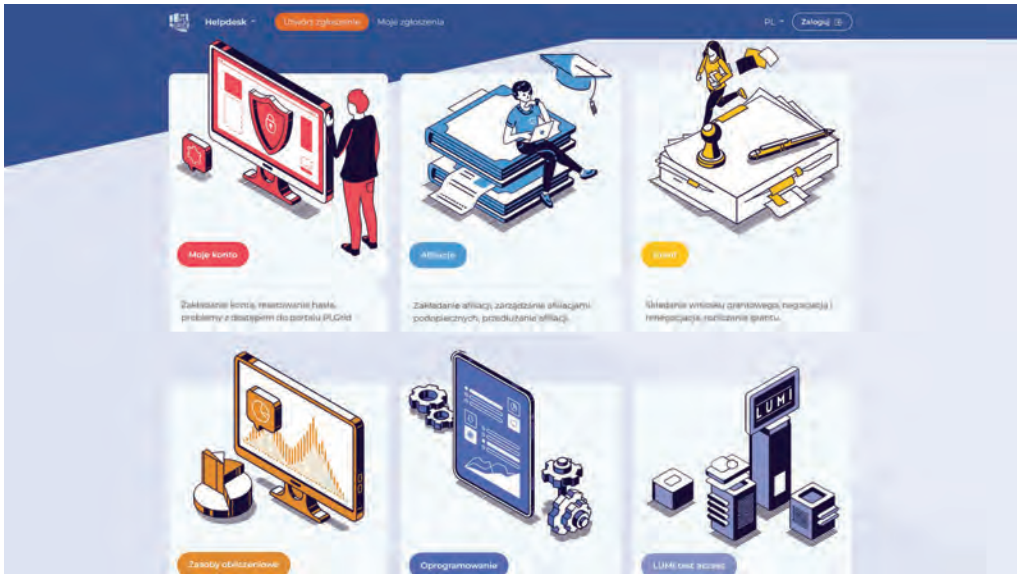
PLGrid Portal – the login screen

The offer for users

Scientists and research teams from Poland, interested in using the PLGrid infrastructure for calculations and large-scale simulations can use the following resources, tools and services free of charge:

- Computing services – the PLGrid infrastructure offers Users the possibility to perform scientific calculations, among others based on computing resources. Among them we distinguish: CPU, GPU and vSMP.
- Data storage services – disk resources are used to store data – permanent or temporary. The resources can be used directly or through dedicated applications and online platforms.
- Specialised packages – collected in the Applications and Services Catalog allow Users to run scientific calculations, among others in the field of biology, quantum chemistry, physics and numerical calculations.

- Graphical interfaces – the PLGrid infrastructure offers a number of tools and services for graphical tasks maintenance and file management, used to organize Users' computational experiments.
- Collaboration tools – a suite of tools (including Confluence, JIRA and Stash) that constitute great support in organizing cooperation within research teams and among Users.
- Help and support – users of PLGrid resources can use the Helpdesk – an effective support system – in solving any difficulties related to functioning of the PLGrid infrastructure.



PLGrid Helpdesk – the login screen

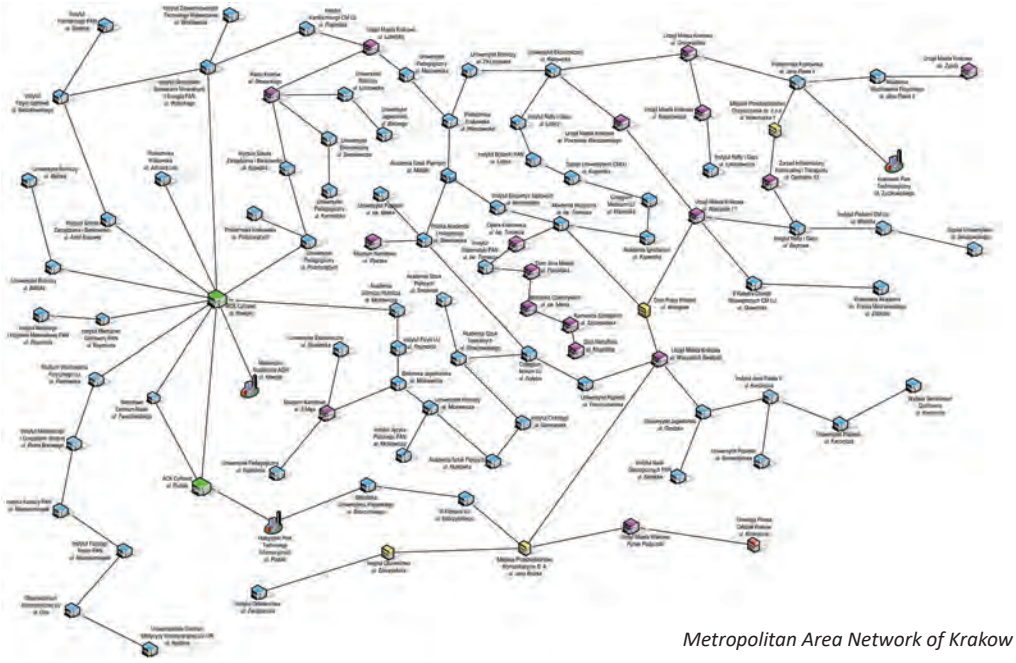
Access to the PLGrid infrastructure allows scientists to scale up calculations carried out in the framework of scientific research and contributes to the expansion of international scientific cooperation.

The PLGrid infrastructure is part of the pan-European infrastructure built under EGI (European Grid Initiative), which aims to integrate national computing infrastructures into one sustainable production infrastructure.

Through the PLGrid infrastructure, scientists can also take advantage of the supercomputer LUMI resources, because as a result of the activities of the PLGrid Consortium aimed at integration of the national computing infrastructure with the LUMI infrastructure, Polish scientists have been provided with powerful computing power of this supercomputer. Access is granted via the PLGrid Portal.

Metropolitan Area Network

One of the major characteristics of the present science is complexity of research challenges, including their multidisciplinary character, use of heterogeneous models, resources and massive amount of data produced by a variety of sources. Research is not performed by a small group of scientists anymore, but by international consortia. To bind those usually geographically distributed resources together, fast and reliable network connectivity is essential. Therefore, one of the principal tasks of the ACC Cyfronet AGH is the development and maintenance of the Metropolitan Area Network (MAN) to achieve its availability 24/7.



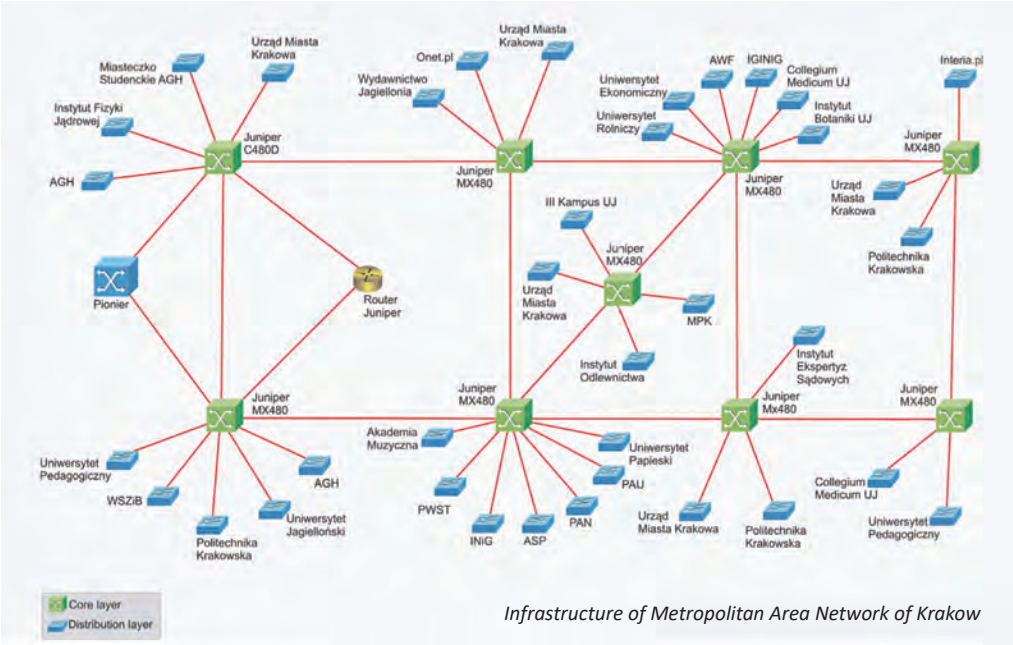
Metropolitan Area Network of Krakow

Main characteristics of MAN

It is impossible to attain high network availability without its continuous development and adjustment to users' needs. The length of dedicated fiber-optic links reached this year over 200 km. The core links of the network are located in the Old Town area and reach the academic campus of AGH University of Science and Technology. Furthermore, the network also covers Bronowice, Krowodrza, Czyżyny and Nowa Huta zones. The recent expansion of the network included such distant research centres like Prokocim, Borek Fałęcki and the 3rd campus of the Jagiellonian University in Pychowice. Development of the core backbone also includes other directions, up to the borders of Kraków.

The fiber-optic infrastructure is the basis of the MAN operation. ACC Cyfronet AGH makes efforts to include in it the largest possible number of university facilities and research institutions. At the same time, due to the ever-growing role of modern communication means, in everyday work, it is crucial that fiber-optic infrastructure, in addition to high bandwidth, could also ensure secure communication. It is realised through the use of backup links, which allow maintaining the continuity of operation when primary routes are broken.

The core data link layers are implemented using top-quality equipment with 1 and 10 Gb Ethernet technologies, while 100 Gb interfaces are gradually being introduced. Each of the backbone network switches is connected with at least two and sometimes even three neighbours for automatic and transparent recovery in case any network device or link fails. Our users can obtain fiber-optic connectivity to the network via 10/100/1000 Mbps or 1 Gbps Ethernet cables as well as through traditional modem uplinks.



Infrastructure of Metropolitan Area Network of Krakow

The Metropolitan Area Network is directly connected to Warsaw, Katowice, Bielsko-Biała and Rzeszów through the PIONIER network. Currently, the links can serve up to 2x10 Gbps capacity. High Performance Computing centres in Poland (Gdańsk, Kraków, Poznań, Warsaw and Wrocław) are integrated with links of 2x100 Gbps capacity. The PIONIER network also enables communication with major national and foreign computing centres. International connectivity is achieved through the GEANT scientific network with 100 Mbps capacity. In addition, the reserve connection with 5 Gbps capacity is established to the Telia Carrier Poland and Lumen Technologies networks.

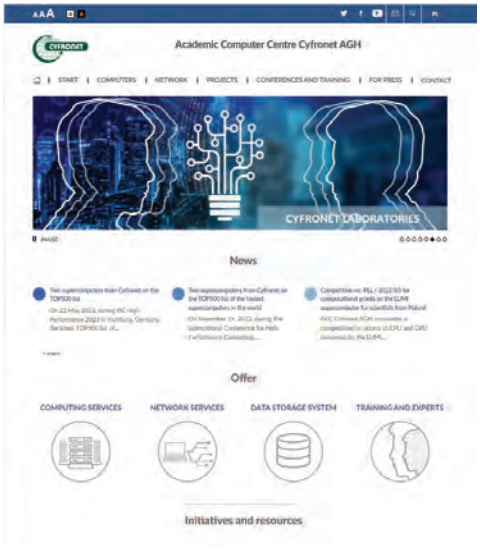
Network services provided to the users

From the beginning of the Polish Internet (mid 1991) ACC Cyfronet AGH has been actively participating in the development of the telecommunications infrastructure and, what is very important, the wide range of Web services. Those include:

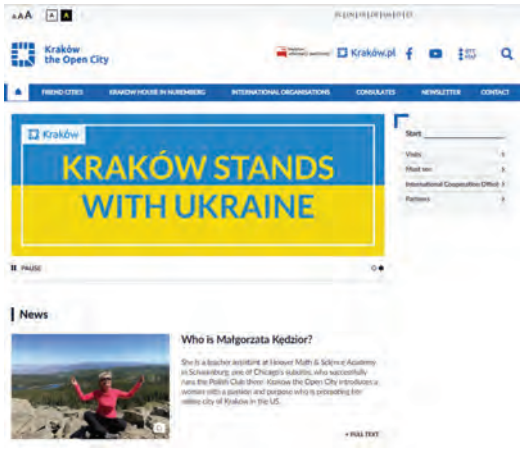
- **e-mail** accessed via SMTP protocol or web interface *http://poczta.cyfronet.pl*,
- **www**: CYFRONET operates a set of web sites, which in addition to news from the world of science, present information on the culture, sights and many other fields,
- **news**: discussion groups covering all areas of interest – from highly specialised scientific to general-purpose boards,
- **dns**: domain name system servers – performing translations of network domain names to IP addresses for users of the Krakow MAN,
- **ftp**: CYFRONET mirrors major international software archives, providing shareware and freeware applications for MS Windows and UNIX systems. The establishment of this service has significantly reduced the traffic on CYFRONET's international links while at the same time enabling faster downloads of software for users of the Krakow MAN,
- **eduroam**: provides the academic network access at all locations on eduroam in the world with a single authorised account,
- **box**: a network drive (*http://box.cyfronet.pl*) allowing file exchange and synchronisation. The drive can also be accessed from mobile devices via a dedicated application.

Network services in numbers in 2022	
Number of e-mails	> 19 000 000
Number of e-mail server sesions	> 53 000 000

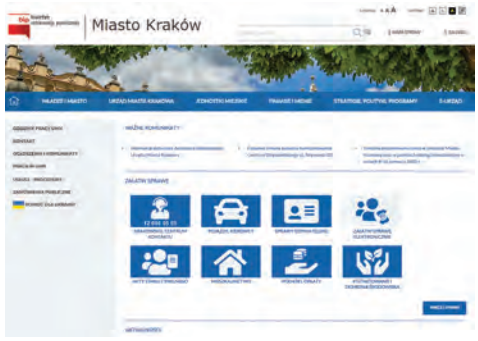
Portals and mobile applications



The Centre does not limit its activities to the scientific areas only – it also contributes to the development of the information society. The Web server at ACC Cyfronet AGH serves as an Internet hub for the entire Kraków scientific community. The Centre continues to develop and extend its Web portal, which has gained substantial popularity over the years. In 2021, a new version of the portal was published, including a version for mobile devices.



Cooperation with Kraków authorities is of particular importance for the Centre. The agreement between the Municipality of Kraków and CYFRONET, regarding the promotion of the City has resulted in the creation of an up-to-date portal. Aside from scientific information the portal introduces its readers to the culture, historic sites, tourism, local transit and many other aspects of life in Kraków.



In collaboration with the City Hall, the Centre has been developing and running the Internet Bulletin for Public Information in the Kraków Region. In 2005 this collaboration was extended in order to provide content services for municipal units, libraries, schools, etc.

In 2007, the "Magical Kraków" web portal – *www.krakow.pl* was nominated for the World Summit Award as the best e-Government service in Poland. The mobile version of the portal was awarded at the conference Mobile Trends, Mobile in 2012 as the best city mobile web site in Poland.

Cooperation with the City Hall explores also the area of mobile devices. CYFRONET has developed – among others – a mobile application "Kraków.pl". The app can be used as a Kraków city guide, a source of important information like phone numbers, info points, consulates or pharmacies. The most important part of this app is the ability to check all those places on an offline map. Our app is available in a few languages.

Computational resources

ACC Cyfronet AGH provides a mature computing infrastructure for Polish science based on five main pillars. Furthermore, complex support and training are available for the users.

Computational resources

Athena, Ares and Prometheus supercomputers provide:
15 PFlops
100 000+ cores
600 GPGPUs
950+ TB RAM.

Storage

45 PB of disk and 30 PB of tape storage space and fast scratch Lustre filesystems enable big data processing and analyses.

Scientific software

Vast portfolio of tools, libraries and scientific applications for research in various fields of science.

Tools for scientific collaboration

Tools and services such as Bitbucket Git repositories server and JIRA issue & project tracking solution ease scientific projects coordination and communication between researchers.

Computational cloud

Cyfronet's PaaS based on OpenStack provides elastic solution for computational environment which can be easily adapted to researchers' needs.

Advanced computing platforms and domain-specific services

Among the scientists conducting research with use of high-performance computers and large storage resources there is a need for different types of interaction with a computer or with the infrastructure. To address these needs Cyfronet provides a number of advanced IT platforms and dedicated services that hide the complexity of the underlying IT infrastructure and, at the same time, provide the functionalities important from the point of view of scientists from the particular field, precisely tailored to their needs.

Together with computing infrastructure we provide a selection of tools, which enable researchers to perform complex, large-scale experiments and manage their results in an easy way. The efficiency of the performed analyses and the safety of their associated data are guaranteed by appropriate IT solutions, benefitting from the extensive experience of Cyfronet's developers. The platforms have been successfully applied in the PLGrid Program for domain specific grids. As we mentioned before we have prepared more than 70 tools, platforms and services gathered into 27 scientific domains dedicated for important scientific topics and strategic fields of Polish science. All those services are provisioned in the framework of the PLGrid infrastructure, allowing Polish scientists and their foreign collaborators to access it in a convenient manner.

Among others, at the Centre we offer advanced tools and graphical interfaces that enable construction of dedicated environments for scientific research, building application portals, conducting virtual experiments, visualization of calculations' results, executing complex scenarios with parallel tasks, as well as supporting uniform and efficient access to data. All of these services are important support for researchers, as they have an impact on improving and, where possible, automating the work of research groups, what greatly accelerates obtaining research results. On subsequent pages we will learn about capabilities of selected services.

Invitation to cooperate

We are looking for people interested in development of domain-specific services. We also offer support in scientific research.

We encourage scientists to send us their program codes for the compilation by the experts at the Centre. After installation, we provide assistance in their effective use. We also enable the use of scientific software licenses held by research groups.

photo: Adam Frączek, KSAF AGH

photo: KSAF AGH

The Cloud Computing in PLGrid

The PLGrid infrastructure has been designed with particular focus on scientists and their needs. Its character allows for easy adaptation to, even sophisticated, research challenges performed by different groups of scientists – from small research teams up to international consortia of researchers. To fulfil their requirements, in addition to computing and storage platforms, we offer the PLGrid Cloud Computing Platform.

– *It has been a standard for many years to offer our users cloud solutions. A user can easily connect to a requested set of virtual machines (VM), with full access rights to the operating system. To achieve high security, all the VMs operate in a dedicated, local area network. Particular services can be accessed from all over the world, easing cooperation between scientists* – says Kazimierz Wiatr, Director of the Center.

There are several advantages of the cloud computing we would focus in particular:

- The Cloud increases elasticity of research, as scientists can independently tune the virtual machines to their specific needs.
- The catalogue of VMs offered by PLGrid contains many OSes. Thanks to this, users can run their software applications with Operating Systems other than Scientific Linux, including Windows or other Linux OSes.
- With Cloud, it is easy to build and put in operation a test environment. This feature is very convenient for scientists developing their own software. Any test task can be then easily performed and its results analysed.
- It is possible to maintain a communication with already executed computing job. In addition, every virtual machine can be easily duplicated, even in thousands of copies or more.
- The Cloud platform is also the best and in many cases the only solution for running jobs with legacy software packages. In a secure LAN environment even old, deprecated operating systems can be used. This feature is also a solution for dispersed international groups using variety of different packages for their research. Every group can run their own computations and easily share their results with others.

Currently users have access to:

- virtual machines (VM instances up to 64 cores and 250 GB RAM with access to fast disk resources),
- OpenStack technology (possibility of easy expansion and adaptation of instance types to the needs),
- shared storage (compatible with S3 and S3 Glacier protocols).

Onedata is a global data management system, which provides transparent access to data stored on distributed storage resource managed by multiple providers. Onedata can scale to meet the needs of small user communities or large federations of users and storage providers, making it a perfect solution for large research initiatives, long-tail of science as well as for commercial purposes. Onedata allows users to rely on a single solution for managing their personal as well as research data sets and access them efficiently on any machine, from personal laptop as well as from a Cloud virtual machine.

Onedata provides a unique federation system based on zones, which enables storage providers to organize into trusted federations and allows users to easily request storage resources from providers within a zone.

Features for users

- Unified access to data stored on heterogeneous storage systems distributed across the infrastructure. With Onedata, users can access their data from anywhere, as the system automatically replicates and transfers necessary blocks on demand.
- All data is organized into *space*, which can be regarded as virtual folders or volumes, accessible from any client machine via POSIX protocol.
- Easy to use web based Graphical User Interface for data access, discovery and management.
- Support for easy data sharing and collaboration with other users, while ensuring security through custom Access Control Lists and creation and management of user groups.

- Open data publishing functionality integrated into the user interface, enabling publication of prepared datasets, registration of DOI identifiers and indexing in open access portals.

Features for administrators

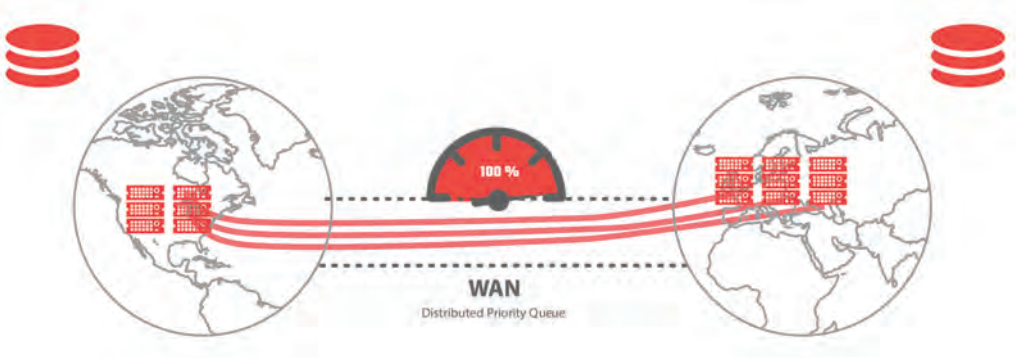
- Simple deployment based on Docker containers using a friendly command line client.
- Easy storage support for user requests based on secure tokens.
- Complex monitoring information available on all aspects of the system, accessible through REST API or directly visualized in the administration panel of the Graphical User Interface.
- Support for multiple storage backends including POSIX based storage (e.g. Lustre), Amazon S3, Ceph, OpenStack SWIFT, and GlusterFS.

Features for developers

- Easy integration with Onedata services using REST API and CDMI protocols.
- Flexible authentication and authorization of requests based on Macaroon tokens.
- Complete reference documentation of the REST API including sample clients for several programming environments.

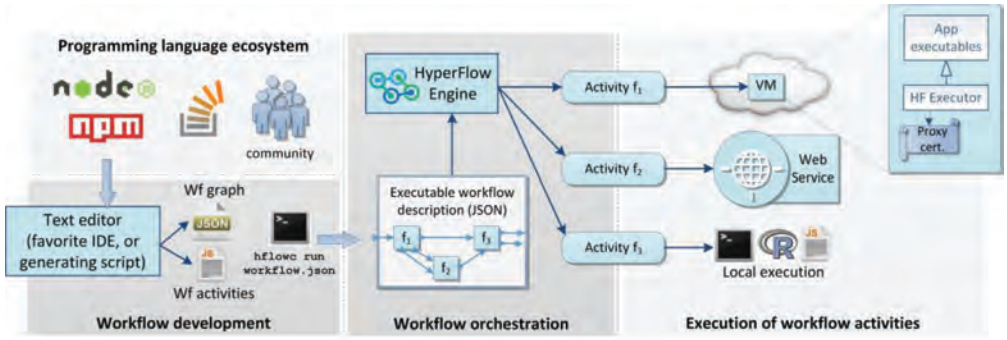
Onedata users

Onedata is currently deployed and evaluated in several initiatives in Europe including Polish National Grid infrastructure PLGrid, INDIGO-DataCloud, EGI DataHub, Human Brain Project and Helix Nebula Science Cloud. In HBP it has proven to meet the users' hard requirements of real-time brain visualization use case.



More information: *https://onedata.org*

HyperFlow is a lightweight tool that enables orchestration of scientific applications into complex pipelines or *scientific workflows*. HyperFlow aids users in composing their applications into workflows, deploying them in the cloud, and executing them.



Workflow programming

A workflow in HyperFlow is described as a graph of its activities (called *processes*) using a simple JSON-based data structure. Workflow activities perform the actual scientific procedures – steps in the scientific pipeline. In HyperFlow, workflow activities can either be implemented in JavaScript or mapped to executable programs. The JavaScript code is executed by the HyperFlow engine in the context of the `Node.js` runtime. An experienced workflow developer can thus take advantage of a mainstream programming ecosystem – large community, advanced tools, thousands of libraries and other resources – instead of using a proprietary development environment. Consequently, workflow activities can easily be programmed to invoke external Web Services, or execute local commands as part of the scientific pipeline defined by the workflow.

In the second option, the workflow developer can choose not to implement any JavaScript code, only associate each workflow activity with a previously prepared Virtual Machine image where appropriate programs are installed, and specify commands that are to be executed when a given workflow activity is triggered.

The availability of these two programming approaches makes HyperFlow equally suitable for experienced programmers / software engineers who desire low-level programming capabilities and high productivity, and domain scientists who are not experts in IT technologies and only wish to construct scientific pipelines out of existing modules.

Workflow deployment

HyperFlow automates workflow deployment in the cloud. The user only needs to prepare a configuration file specifying the mapping of workflow activities onto available Virtual Machine images, while the HyperFlow tool takes care of the rest. The user invokes a simple command *hflowc setup* which results in creation of appropriate VM instances in the cloud. These VM instances contain the workflow runtime environment and the scientific applications invoked from the workflow.

Workflow execution

After the workflow instance has been created in the cloud, the user executes the workflow simply by invoking *hflowc run <workflow_directory>*. Every workflow runs with its own instance of the HyperFlow runtime environment. Consequently, different workflow runs are isolated from each other which increases security and reliability.

The HyperFlow cloud runtime environment (called *HyperFlow Executor*) automatically takes care of transferring input data from the user directory to Virtual Machine instances, invokes the application executables and uploads output data back to the user directory. A variety of data transfer options are available, including a network file system, secure gridftp, and Amazon S3.

Applications

HyperFlow has become a part of several larger systems where it has been used for a number of applications. In the PLGrid infrastructure, HyperFlow serves as a workflow management system that enables the users to run scientific workflows in the cloud. An example application is a workflow-based solver for finite element meshes which can be applied to diverse problems. HyperFlow is also being integrated with the PaaSage middleware (*http://www.paasage.eu*) as an execution engine for scientific applications deployed in a multi-cloud environment. In the ISMOP project (*http://www.ismop.edu.pl*), HyperFlow is a component of a flood decision support system used to orchestrate flood threat assessment workflows. Hyperflow will also be used in parametric computing and workflow processing, as a replacement for Scalarm technology.

Contact

HyperFlow is developed and maintained by the DICE team (*http://dice.cyfronet.pl*). Please feel free to contact us in case of any questions or suggestions.

HyperFlow source code and manuals are available at *https://github.com/hyperflow-wms/hyperflow*.

Model Execution Environment

The Model Execution Environment (MEE) is a software stack which facilitates the execution of computational workflows on high-performance computing infrastructures, including those available at ACC Cyfronet AGH. The platform's goal is to ensure that computations can be executed straightforwardly by domain scientists, i.e. researchers who do not possess intimate knowledge of the specifics of interaction with computing clusters and other large-scale computing systems.

Pipelines, steps, and models

Within MEE, computational workflows are represented by the so-called pipelines, i.e., collections of computations (each of which is called a step) where the outcome of one computation provides input for another. MEE provides a wide range of facilities enabling users to design steps, arrange them into pipelines, and execute these pipelines on the available computational resources.

Each step is based upon a collection of computational artefacts (executable code) stored in a GitLab repository. This is referred to as the model. When a step is called for execution, MEE automatically uploads the requested model to the HPC infrastructure and monitors its performance on the input data provided. Users can select a specific version of the given model when launching the pipeline (based on Git versioning mechanisms), thus facilitating traceability and repeatability of computations.

In addition, pipelines can be executed in either automatic or manual mode. An automated pipeline will be executed in its entirety, while a manual pipeline contains a breakpoint at the end of each step, asking the user to manually request the processing of any subsequent steps. This enables users to

download and review interim results and potentially cancel the execution of pipelines which are not expected to yield useful output, thus preserving computational resources.

All MEE features can be accessed via a user-friendly web-based UI. Furthermore, MEE provides programmatic access, which enables integration with higher-level software tools.

Research data management

As MEE schedules and monitors the execution of computational pipelines on HPC resources, care must be taken to manage the associated research data, ensuring that the appropriate input is made available to the underlying models, and that results can be retrieved from the HPC infrastructure. To this end, MEE provides a set of data management interfaces where users of the infrastructure can upload input files and download results. The platform itself manages HPC data storage resources and provides automatic stage-in and stage-out capabilities for research data, along with a set of top-level UI interfaces for its users.

Security

The Model Execution Environment is integrated with PLGrid authentication and authorisation mechanisms. All users of the PLGrid infrastructure can use their login to authenticate themselves with MEE and subsequently schedule and run computations using their PLGrid accounts. Moreover, MEE makes use of PLGrid computational grants assigned to researchers. Pending computations are executed in the context of specific computational grants, which can be predefined within the platform.

Organisations

Externally, MEE provides a set of distinct workspaces dedicated to individual research teams. These are referred to as organisations. Each organisation has a distinct entry point to MEE (i.e., a distinct URL) and can define its own pipelines and pipeline steps, as well as manage its own set of research data. MEE implements compartmentalisation, where each organisation can be managed separately, providing access to a distinct group of users.

Applications

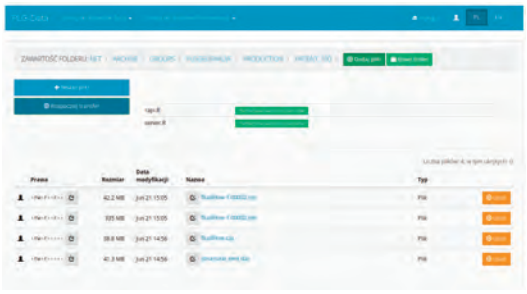
To date, the Model Execution Environment has been applied in multiple research projects and provides services to various research groups. The list includes the EurValve, PRIMAGE, and InSilicoWorld European projects, the POLVAS consortium, and the Sano Centre for Computational Medicine, along with several ad-hoc research collaborations for which individual MEE organisations have been defined.

https://mee.cyfronet.pl

Simple tool for file management on a computing cluster

PLG-Data

PLG-Data is a tool for the management of data stored in the PLGrid infrastructure. It comes with a user-friendly web interface and allows to upload, download, browse, delete and rename files and folders. It also helps with the management of access rights for members of a research group or external collaborators. It is currently integrated with Prometheus, Ares and Athena (through Ares's file system) supercomputers, and in the future, it may be integrated with new Cyfronet supercomputers.



The set of functionality built into the tool includes, among others, the following:

- downloading files from the cluster to disk,
- adding new files and folders, and removing existing ones,
- renaming files or folders, and changing access rights to them,
- quick navigation to home, scratch and group folders through a handy pull-down menu,
- easy preview of image files without downloading them to a local disk.

Thanks to a specific construction of the URL address to particular files, the tool enables easy sharing of file location with other people, for instance, through copying the browser address bar's content to an e-mail message or an IM communicator. The receiving person will be able to download a file or view the contents of a folder with one click – as long as that person is allowed access to the specific resource.

The service is secured with encrypted HTTPS protocol (between the user's computer and the PLG-Data portal) and the specialised GridFTP protocol (between the service's portal and the computing cluster). The application of such techniques allows the user to manage their files securely. A person who uses PLG-Data does not receive any additional access rights to files stored in the computing cluster, apart from the rights that the person already has.

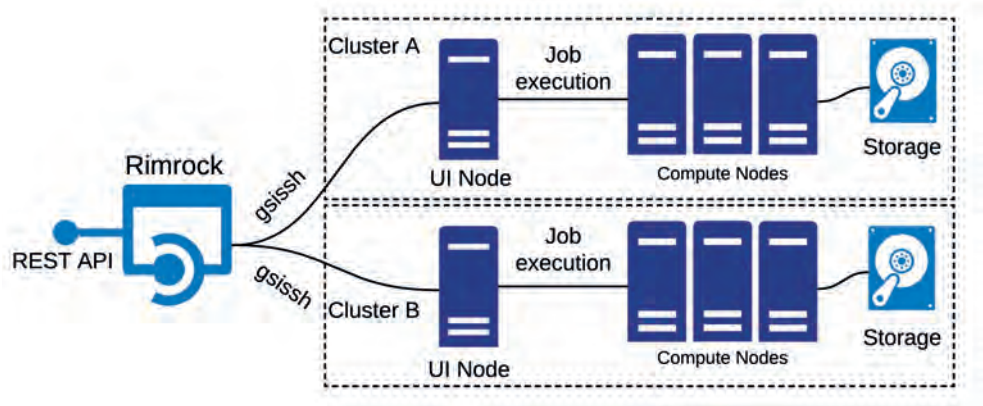
Logging in to the tool is done using either the PLGrid user-password pair, or a p12 certificate installed in one's browser. The service is available either in Polish or English. An advanced programmer's interface (API) helps developers to integrate their platforms, tools and services with the file storage inside the PLGrid infrastructure.

PLG-Data service address is: *https://data.plgrid.pl*

Rimrock, one of the services of the PLGrid infrastructure, enables the management of scientific computation and result handling with the use of modern interfaces based on REST (Representational State Transfer). REST is a well-established programming pattern often used in applications with distributed architectures. Access to services, applications and advanced scripts deployed on the infrastructure becomes straightforward using REST.

Readiness for various applications

Applying REST principles in implementing the rimrock service allows one to use its functionalities independently of any programming language. It is, therefore, possible to create web and desktop applications as well as prepare advanced computation scripts (e.g. with the use of *Bash* and the *curl* command). An interesting approach also supported by the service is the ability to develop web applications, which can be run solely in the user web browser, minimising the role of server-side software.



Support for several job management systems

The rimrock service uses the Slurm job management system, which ensures support for its unique features. It allows for easy integration of legacy applications in newly developed systems. Access to computation results is facilitated by hiding the internal file transfer protocol (*GridFTP)* and by grouping the results according to the executed jobs.

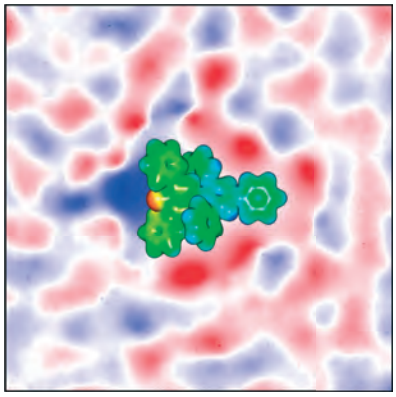
Data security

Data exchanged with the rimrock service is transferred with secure HTTPS connections, and for user authorisation a temporary user certificate (so-called *proxy*) is used.

https://submit.plgrid.pl

Chemistry and Biology – electronic structure and molecular dynamics software

Modern computational chemistry requires constantly increasing resources. More and more computational power is needed to make large systems (especially current challenges of nanotechnology or biological sciences) tractable and improve the accuracy of obtained results. Fortunately, constant progress in computer technology and specialised software offered by Cyfronet meet this demand and enable various chemical computations.



A. Eilmes, P. Kubisiak: Electrostatic potential of an ionic liquid around the solvated dye molecule

Cyfronet clusters' nodes provide up to 1.5 TB of RAM and 48 cores per physical node, which enables quantum chemical computations that require a large amount of memory or a high number of cores with shared memory. Moreover, the fast InfiniBand interface allows good speed-up of calculations if distributed over many nodes. Various quantum chemistry codes also need fast and broad I/O for storage systems. The parallel-distributed Lustre scratch file system and the possibility to use RAMDisk on selected nodes enable that.

Efficient quantum chemistry computations also rely on efficient installation of scientific software and its proper usage. Our administrators' team has the necessary skills, knowledge and experience in installing various applications and efficient running computations. Our portfolio of software used in chemistry contains many packages. Among them, there are:

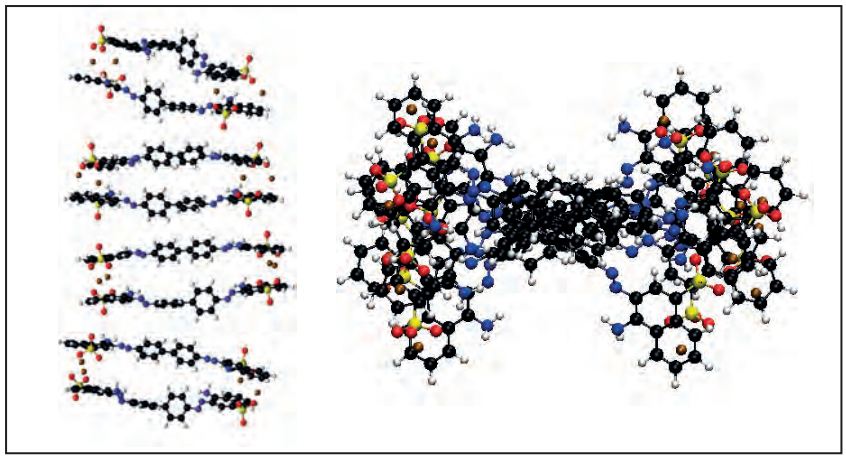
- Versatile and widespread used quantum chemistry codes such as **Gaussian**, **GAMESS US**, **NWChem**, **Schrödinger**, **Q-Chem, Psi4, ORCA** and **TURBOMOLE**, which are capable of calculating electronic structure and various properties of diverse molecular systems using both *ab initio*, density functional theory and semi-empirical methods.

- **Molpro**, **CFOUR** and **Dalton** suites to analyse chemical systems with great accuracy using sophisticated methods such as CC (up to CCSD(T)) and MCSCF.

- Amsterdam Modeling Suite **(AMS, DFTB, MOPAC, COSMO-RS)** provides methods to examine various properties (especially spectroscopic, such as NMR and ESR spectra) of molecular systems with reliable relativistic ZORA approach, COSMO-RS method and all-electron basis sets for the whole periodic table. With addition of versatile and well-constructed GUI of AMS (**AMSInput**, **AMSViev**, etc.) ADF package is used by many of our users.

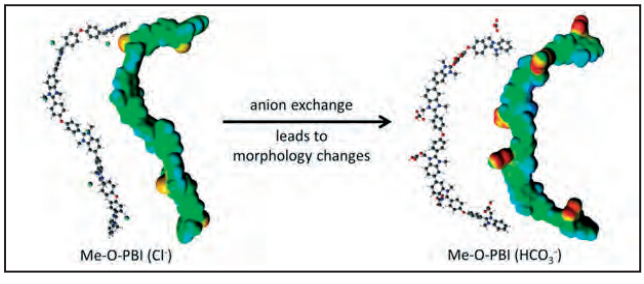
- Several packages, which could be used for solid-state systems. Among them **BAND**, **Quantum ESPRESSO** and **SIESTA** are worth mentioning.

- **AlphaFold** using machine learning to analyze the geometric structures of proteins.
- **Desmond**, **Gromacs**, **Amber**, **LAMMPS**, **NAMD**, **Tinker-HP**, **CPMD**, **CP2K** and **Terachem** suites for molecular mechanics and molecular dynamics simulations of systems containing hundreds of thousands and more atoms.



O. Klimas: Optimized stack of eight Congo Red molecules seen from different perspectives

Nowadays general-purpose computing on graphics processing units (**GPGPU**s) in many scientific domains provides great speed-up of calculations (up to several orders of magnitude). In our computing Centre, some of nodes provide possibility of such calculations on **CUDA** enabled **GPGPU**s (up to eight cards per node). Among software prepared to run on graphical processors our administrators' team prepared quantum chemical packages such as **GAMESS**, **Terachem**, **NAMD**, and **Quantum ESPRESSO, Tinker-HP**. Our experts extensively collaborate with several, mentioned above, packages developer teams. The Cyfronet team prepares and helps with adjusting the dedicated computing environment for our users.



Electrostatic potential of molecules in anion exchange membrane. Published by W. Germer, J. Leppin, C. Kirchner, H. Cho, H. Kim, D. Henkensmeier, K. Lee, M. Brela, A. Michalak and A. Dyck in Macromol. Mater. Eng. 2015, 300, 497–509

Machine learning (ML) and artificial intelligence (AI)

AI-accelerated data analysis is making great strides in many research domains, including materials as well as life science, linguistics and social science. The ability of neural networks to learn from complex data may significantly improve data analysis, classification and pattern detection, with potential applications in many systems, including image recognition, language processing and optimisation.

The Cyfronet supercomputing centre faces up to these challenges and prepares several packages:

PyTorch is a package, specifically a machine learning library for the Python programming language, based on the Torch library. It enables implementation of complex Deep Learning algorithms from the Natural Language Processing, video and images processing and many other areas. It can be used for modeling new architectures in the field of machine learning with focus on experiments.

TensorFlow allows, like Pytorch, to implement models based on the tensor flow paradigm. Due to its character and static representation graph, it allows for efficient optimisation of models training and inferences with respect to the computing platform.

Keras is a library used for designing neural models. It is an external API for engines based on TensorFlow, Microsoft Cognitive Toolkit, Theano, or PlaidML. It has been designed to enable fast experimentation with deep neural networks. It focuses on being user-friendly, modular, and extensible.

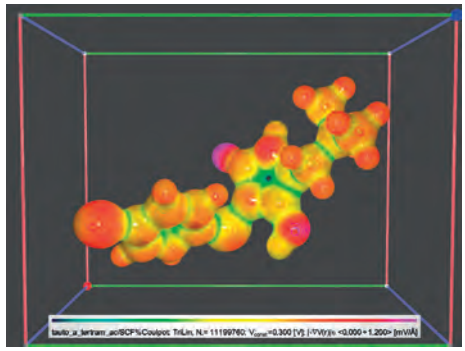
Scikit-learn is a software machine learning library for the Python programming language. It features various classification, regression and clustering algorithms including support vector machines, random forests, gradient boosting, k-means and DBSCAN. It has been designed to interoperate with the Python numerical and scientific libraries NumPy and SciPy.

SchNet is a deep learning architecture that allows for spatially and chemically resolved insights into quantum-mechanical observables of atomistic systems.

Horovod is a distributed training framework for TensorFlow, Keras, PyTorch, and MXNet. The main goal of Horovod is to make distributed Deep Learning fast and easy to use.

Data Visualization, POVRay/ScPovPlot3D

Data visualisation enables analysis and understanding of the results of even very complex numerical calculations, especially multidimensional or time-dependent. Most applications for numerical calculations have a module that generates their visualisation. Python has a matplotlib or VTK+ module, while Matlab or R also have graphic libraries. The situation is similar with regard to geovisualisation programs (GIS) or chemical calculation programs. Unfortunately, no matter how much these programs are refined, the result of their operation is limited by the Cartesian product of available (and compatible) options.



Janusz Opiła: Electrostatic field configuration around the polymer molecule. An equipotential surface with a trilinear approximation is shown, color encodes the electric field intensity module

Overcoming of this limitation, at least for the purpose of creating a prototype of visualisation style for later implementation in a dedicated package, is possible, but requires using a general purpose graphics program, for example 3DMax, Blender or POVRay. However, only the latter is equipped with a scripting language (*Scene Description Language* – SDL), which allows for programmatic, non-interactive creation of visualisations, so is useful for mainframes. As the use of countless SDL language options requires quite persistent studies, a dedicated API was written in the form of a set of specialised modules named the "ScPovPlot3D". This is not a completed project as further extensions are still being added. Thus it may be called a beta version, but mature and working. Currently, the project is in version 4.0 and is hosted on GitHub (URL: *https://github.com/JustJanush/Plot3Dv4*) – the multiplatform API requires POVRay at least in version 3.7.

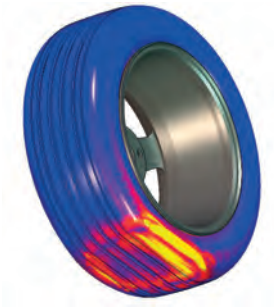
The most important modules are:

- VectorField.inc – hybrid vector field visualisation using widgets and/or field stream tubes,
- Potential.inc – hybrid visualisation of scalar fields, on regular and irregular meshes with trilinear or centripetal Catmull-Rom cubic interpolation,
- BPatchSurf.inc – hybrid surface visualisation based on data on regular or irregular grids with implemented simple kriging (KDE),
- Mesh2Surf.inc – hybrid visualisation of data defined on regular 2D grids (z=f (x, y)),
- TextExt.inc – extended 3D text formatting, oriented to the presentation of mathematical formulas.

Janusz Opiła: Terrain visualisation based on altitude data collected on an irregular grid and textures obtained from the Google Earth Pro application vicinity of Karlobag, Croatia). Own study: DOI: 10.23919/MIPRO.2018.8400037

If necessary, the package's developer provides technical support. Contact information: *https://skos.agh.edu.pl/osoba/janusz-opila-2390.html*

CAD/CAE applications



Computer-Aided Design and Computer-Aided Engineering applications are essential tools in developing and building almost everything – from car parts to buildings. Through computer simulations, engineers can check the durability of constructs and devices; perform linear and non-linear structural analyses of contact phenomena, plasticity, recoil, etc. CAD/CAE software provides analysis of thermal conductivity, radiation and phase shifts. Significant for science are also fluids simulations: velocity fields, pressure fields, heat distribution, chemical reactions, etc.

Cyfronet's users can resolve all these tasks thanks to CAD/CAE packages of ANSYS, ABAQUS, FLUENT, MARC and OPERA.

ANSYS is a complex structural simulation package with an intuitive graphical user interface, supporting scientists from nearly any area of science or business. Results are calculated with high precision and may be presented by plots or tables, for example, isosurface diagrams and deformations. Computational capabilities of ANSYS are very high and involve: harmonic and spectral analysis, statistics and dynamics.

ABAQUS is devoted to solving problems in the industry using finite-elements analysis. A user can prepare a combination of finite-elements, materials, procedures of analysis and sequences of loads, according to individual requirements, to simulate vehicle loads, dynamic vibrations, multibody systems, impacts, crashes and much more.

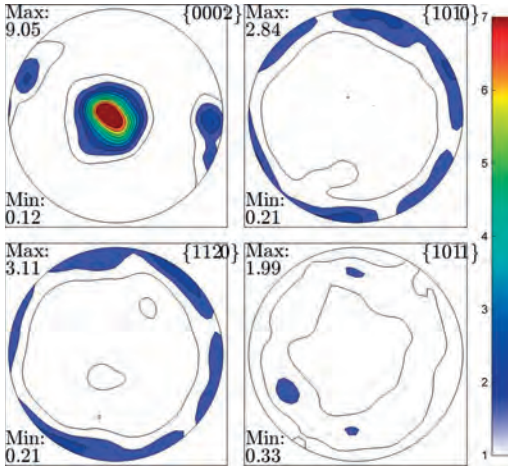
FLUENT software offers the broad physical modeling capabilities needed to model flow, turbulence, heat transfer and reactions for industrial applications ranging from air or liquid flow to semiconductor manufacturing. FLUENT can be used in numerous science domains, including chemistry, metallurgy, biomedicine, electronics, material design and many others.

MARC is a general-purpose, non-linear finite element analysis solution to accurately simulate the product behavior under static, dynamic and multi-physics loading scenarios. It can simulate all kinds of non-linearities, namely geometric, material and boundary condition non-linearity, including contact. It is also the solution that has robust manufacturing simulation and product testing simulation capabilities, with the ability to predict damage, failure and crack propagation. All that can be combined with its multi-physics capabilities that helps couple thermal, electrical, magnetic and structural analyses.

OPERA is a finite element software suite for design and optimisation of electromagnetic devices in 2D/3D. It gives accurate numerical solutions for problems from multiple areas of science, including electrostatics, magnetostatics, low and high frequency electromagnetics. The software gives an ability to design and optimise many types of electrical devices: transformers, motors, switches, micro-machines, MRI scanners and X-ray tubes. It is a powerful virtual prototyping facility to accelerate the design process.

Symbolic math applications

Mathematical applications enable to conduct in reasonable amount of time even very complex and complicated calculations. Users of ACC Cyfronet AGH have access to software that supports calculations in the field of algebra, analysis, combinatorial math, statistics, theory of numbers, geometry or other math areas. Running calculations like integration, differentiation, symbolic processing, matrix operations, approximation and interpolation, Fourier and Laplace Transforms, digital signal processing, etc. is a lot easier. Results can be visualised with appropriate tools. Some of the applications can create interactive 2D and 3D plots. In scientific work, preparation of precise model that most accurately describes analysed issues, is essential.

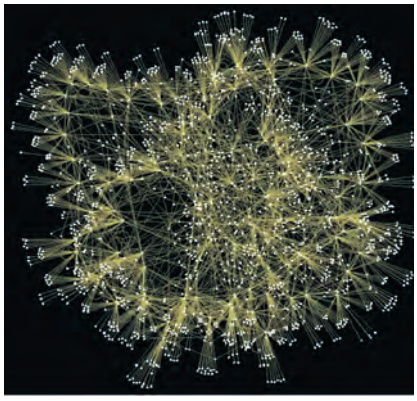


Bartosz Sułkowski: Results of texture simulations by visco-plastic self-consistent model of Zn after hydrostatic extrusion at 250 °C

A good example of software environment, which can be applied in above-mentioned issues, is **MATLAB**. Its modules (Toolboxes) allow performing computations in the field of financial modelling, partial differential equations, linear and non-linear optimisation and much more. It is also possible to use Simulink – the environment oriented for simulations and visualisations from blocks, without the need for traditional programming.

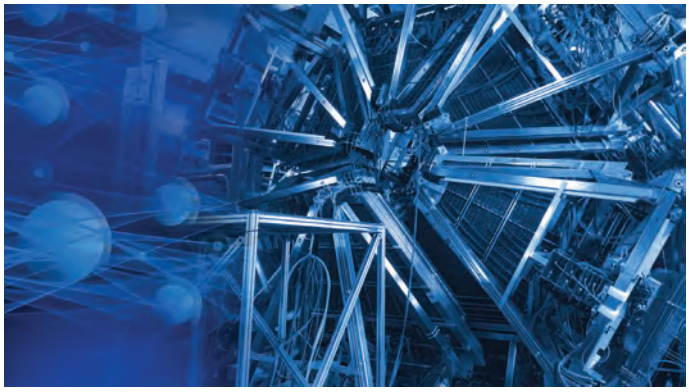
Apart of that environment, users can find in our software a useful application, **MATHEMATICA**, which allows parallel computations with defined precision, dedicated for symbolical and numerical calculations. An advantage of MATHEMATICA is, among other things, a tool for fixing mistakes.

Another example of universal and interactive mathematical software is **MAPLE.** It can be used for simplification of expressions and symbolic processing. It offers databases, enables code generation in other programming languages, creating slideshows with user commands and communication with MATLAB and CAD systems.



Rafał Rak: One minute price returns network for KGHM (the Polish stock company)

Dynamically developing scientific research requires more and more advanced tools nowadays. Among them, IT tools play a huge role, supporting the effective research from the moment of its design to the development of results. Cyfronet, by following the latest solutions and creating its own studies, tries to fulfill an important area of its mission to support science. Dedicated laboratories were established for these needs.



Laboratory of Quantum Computing

The Laboratory was established to conduct research on the use of quantum computers in calculations and to support classical calculations with quantum accelerators.

One of the key tasks is to follow the development of quantum computing technologies and available quantum accelerator platforms in order to use them in dedicated services offered by Cyfronet. We cooperate with other research entities and industrial partners, both as consortium members of joint initiatives and carrying out commissioned works.

Based on our own competences and the exchange of expert knowledge with a network of partners, our team is working on solving the problems that prevent us from wider and more effective using the quantum accelerators in calculations for the benefit of science and economy.

We also act to popularize calculations using quantum accelerators and provide substantive user support. In this regard, we prepare the necessary documentation and materials, conduct training and publish the results of research work.

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

Laboratory of Parallel Algorithms

The Laboratory focuses on the broadly understood computational aspect of parallel algorithms, with particular emphasis on machine learning algorithms, and the possibility of their effective use on large computing clusters.

Due to the increasing amount of data available every day, traditional machine learning algorithms are becoming insufficient and the serial processing paradigm is computationally inefficient. In order to meet the new challenges related to the growing amount of data, in many cases it is necessary to use large computing clusters, which forces the adaptation of the algorithms used to work in parallel mode. In the Laboratory of Parallel Algorithms, we consider both theoretical and practical aspects related to this task. In particular, we focus on the following areas: computer vision, tensor computing, deep networks, low quality image processing, underwater image recognition, hyperspectral data classification, histopathological data classification, data unbalance.

Within the considered domains, we develop new algorithms that use parallel computations, in particular, we study the theoretical and practical aspects related to this phenomenon.

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

Laboratory of Information Methods in Medicine

The main tasks of the Laboratory focus on two spheres. The first is research activity, which includes a thorough analysis and verification of available and potential answers to the challenges found at the border of medicine and information technology. The second one covers the design, development and subsequent operation of dedicated applications and platforms for medical applications. This scope also covers the monitoring of the security status of the developed software as well as data storage and processing mechanisms.

Thanks to the comprehensive approach to the processes: from the identification of the research problem, through the analysis of users' needs, to the final implementation, the Laboratory effectively implements its mission to support the scientific and medical community. As part of the dissemination of expert knowledge, members of the Team publish research results in scientific journals, participate in the preparation of information materials and conduct consultations for users.

Laboratory employees establish cooperation with renowned domestic and foreign research institutes and medical IT centers. The effects of this cooperation are, among others, ongoing and already implemented projects with significant participation of Team members:

- Sano: Centre for New Methods in Computational Diagnostics and Personalised Therapy,
- PRIMAGE: PRedictive In-silico Multiscale Analytics to support cancer personalized diaGnosis and prognosis, Empowered by imaging biomarkers,

- Virolab: A Virtual Laboratory for Decision Support in Viral Disease Treatment,
- Gliomed: Diagnostics of gliomas based on the slowly circulating DNA of the tumor,
- Eurvalve: Personalised Decision Support for Heart Valve Disease,
- CECM: A Centre for New Methods in Computational Diagnostics and Personalised Therapy.

The previous activity of the current Laboratory team is presented in detail on the following website: *http://dice.cyfronet.pl*.

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

Laboratory of Data Processing

The Laboratory designs and implements dedicated applications and software platforms for applications in various fields of science. The Laboratory consists of specialists in the field of software architectures, Front-end and Back-end programming, user interface and user experience design, DevOps, testing, and requirements analytics. The team specializes mainly in:

- development of innovative methods of acquiring knowledge from available data,
- development of technologies supporting open data processing,
- integration of data and knowledge processing systems with existing repositories and e-infrastructures.

The Laboratory establishes cooperation with renowned scientific and research units as part of Polish and international projects. The effects of cooperation include:

- **Construction of the Sat4Envi Portal** (*https://dane.sat4envi.imgw.pl*), providing satellite data from the Copernicus program. The portal enables searching, viewing, ordering and downloading satellite data and their derivative products using only a web browser.
- **Development and maintenance of the EOSC Portal** (*https://eosc-portal.eu/*) as part of a series of projects related to the European Open Science Cloud (EOSC). The portal provides access to the resources of many European e-infrastructures and research infrastructures through a unified user authentication system. EOSC activities focus on the implementation of the Open Science paradigm.

The Laboratory team was also involved in the creation of the **PLGrid Portal** (*https://portal.plgrid.pl*), which provides scientists with many software packages, libraries and scientific tools. In 2022, the new Laboratory of Interdisciplinary Scientific Computing was separated from the Laboratory of Data Processing.

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

Laboratory of Interdisciplinary Scientific Computing

The Laboratory conducts research and development work on the processes of conducting and supporting scientific calculations and the organization of scientific data. The Science Gateways portals elaborated by the Laboratory employees create the possibility of establishing cooperation with external entities: Polish and foreign.

The Laboratory tasks are also focused on:

- implementation of research grants and industrial orders,
- using expert knowledge to solve problems requiring the use of various computing resources, e.g. machine learning technology,
- popularisation of calculations using the tools created by the Laboratory among users,
- essential support for users, realized among others by monitoring needs, developing documentation, and conducting training.

The team of the Laboratory, previously co-creating the Laboratory of Data Processing, developed the proprietary InSilicoLab programming environment, which includes a set of advanced tools and programming libraries that allow for the construction and development of dedicated research portals. Portals based on InSilicoLab are designed in such a way as to gather in one place all the tools that researchers need for *in silico* calculations. The main advantages are:

- easy running of the user experiments, even if they are complex, long and require many calculations,
- the ability to conveniently describe, categorise and search for input or output data.

The InSilicoLab technology is distinguished by striving for the greatest possible usability of the tools built with the help of the environment. This sphere includes both the usefulness for solving scientific problems in a given field, as well as the user-friendliness of the portal for its end-user.

The effects of the Laboratory team cooperation with renowned scientific and research units as part of Polish and international projects include the development of the **EPISODES Platform** (*https://tcs.ah-epos.eu/)* as part of a series of projects related to the European Plate Observing System - EPOS (*https://www.epos-eu.org/*). The portal and the tools organized around it are focused on the study and analysis of seismicity and other phenomena caused by human activity (e.g. exploitation of resources within a mine, creation of artificial water reservoirs). The portal is integrated with the European EPOS infrastructure.

Contact: Joanna Kocot, j.kocot [at] cyfronet.pl

Laboratory of Cloud Technologies

The Laboratory deals with the design and operation of the cloud for science, as well as tools for its effective use. The team is developing comprehensive environments for access to distributed data, taking into account both the issues of secure data storage and processing in the cloud, as well as convenient access interfaces (portals, applications) for the end user.

Bearing in mind the dynamic development of new technologies for processing and storing data in the cloud, the Laboratory constantly conducts research and publishes the results. Using the team's expert knowledge, it actively supports scientific initiatives, including international projects and e-infrastructures.

The flagship product of the Laboratory is **Onedata**: a globally scalable data management system, unifying access to data stored in distributed systems. Onedata responds equally well to the needs of both small user groups and large international research communities. The system enables users to use a homogeneous data management system for both personal and work-related data storage, such as research results, and enables accessing it efficiently from any device.

More information at: *https://onedata.org*.

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

Laboratory of Applications of Computational Techniques

The key aspect of the Laboratory's operation is the development of numerical models and computer simulations based on supercomputer architectures, with particular emphasis on practical use in industry. The Laboratory conducts its own research on computational techniques, as well as tracks and analyses new solutions in their application.

The Laboratory team establishes cooperation within Polish and international consortia, and together with partners it implements projects aimed at developing or improving the production processes of metal products. Among the most important effects of these activities are:

- **VirtROLL - Virtual strip rolling mill**: the main goal of the project was to create a computer system supporting the flexible design of rolling technology for flat products based on the calculation results of dedicated numerical simulation modules.
- **PROTEUS-RS - Long products manufacturing processes optimization strategies to improve the finished product quality by minimizing residual stresses**: the project provides for the production of models and a series of numerical simulations based on high performance computer architectures for the design of the long metal fabrication process.

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

Laboratory of Visual Techniques

The main task of the Laboratory is to produce materials documenting and promoting the achievements of ACC Cyfronet AGH.

The popular science and training films and TV programs that we make show the results of research work carried out by the users of Cyfronet's computing infrastructure.

The Laboratory hosts workshops on journalism, the art of cinematography, sound production, editing, computer graphics and television studio operation. It is also a place of research in vision technologies, acoustics and online transmission techniques.

The Laboratory has film and television equipment for image, light and sound production in studio and outdoor productions.

It also cooperates with professional creators and producers as well as students of Krakow's universities.

The films made by Laboratory's team can be viewed, among others:

- on the YouTube channel of ACC Cyfronet AGH (*https://www.youtube.com/user/CyfronetAGH/videos*),
- on the Pionier.tv platform (*https://pionier.tv/wideo-tag/cyfronet/*).

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

Laboratory of Artificial Intelligence

Our mission

Thanks to our expertise in the field of Artificial Intelligence (AI) algorithms, and our knowledge of HPC computational methods, we can support the scientific community in their AI-based research. Our knowledge and experience allows us to implement machine learning algorithms and dedicated to neural networks effectively using the Athena supercomputer and the AI-dedicated partition of Prometheus' supercomputer available at ACC Cyfronet AGH. Athena is enhanced with 384 NVIDIA A100 GPGPU cards, and the Prometheus AI computing partition is built on four powerful servers equipped with eight nVidia Volta V100 GPGPU cards each. The total computing power of Athena and the partition is approximately 64 PetaFlops, which is over sixty-four quadrillion ($4x10^{15}$) AI calculations per second.

Domain specialisations

Selected applications of Artificial Intelligence in natural language processing (NLP), image processing and time series analysis for various research problems are presented below.

Natural language processing

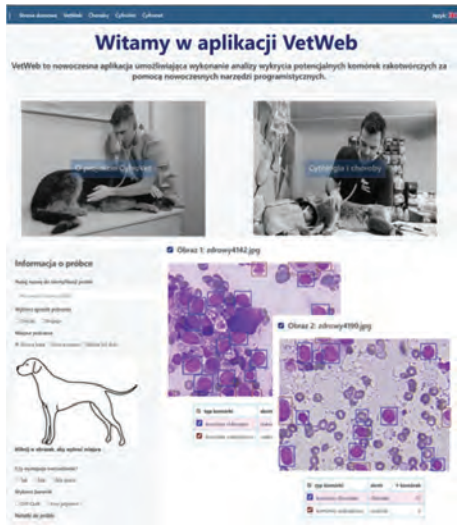
The Laboratory is proud of many years of experience in natural language processing. For example, we have created and developed a tool that allows users to search, compare and classify text documents. The result of this work is the web service called Scholar that is available at ACC Cyfronet AGH on the PLGrid platform. One of the important research problems is the analysis of the impact of the methods for reducing the accuracy of textual data representation on the effectiveness of the NLP algorithms. We have managed to develop alterations of the methods that allow for a 10-fold reduction in computing energy consumption, if compared to the original implementation, with no significant loss of accuracy.

The emergence of neural network-based solutions has revolutionised the NLP field. We research the compression and hardware implementation of the sentiment assessment network, which showed that it is possible to reduce the accuracy of the network coefficients precision to 8 or even 4 bits while maintaining the network efficiency almost unchanged. Additionally, we examine the area of semi-supervised learning, where the amount of available tagged data is very limited, and the output categories changed during the system's operation. Our research has shown that it is possible to develop a solution with an accuracy of up to 98.9%.

Image processing

Neural networks-based image processing is the main focus of the Laboratory interests. The efforts are related to the recognition and detection of objects for the needs of medical applications. The

Laboratory's basic research concerns are understanding veterinary images, which is carried out at Cyfronet as part of the CyfroVet project. An interdisciplinary team of veterinarians and IT specialists working in CyfroVet creates a synergy of knowledge from veterinary medicine, artificial intelligence and cloud software.

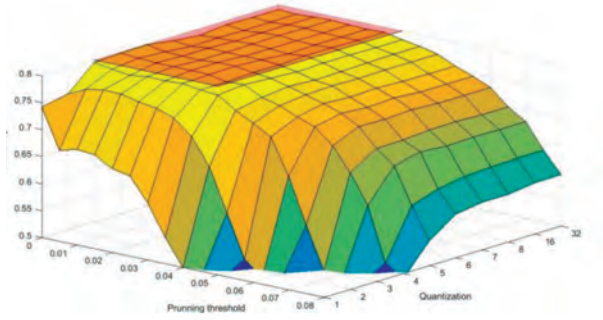


A system for classifying neoplastic lesions has been developed. It is based on images from cytological examination of the skin of dogs. The proposed system has achieved high efficiency for selected most common cancers. In the work, the models of recurrent (resnet) deep neural networks and transformer were used. The practical goal is to use the results of our research in the daily work of veterinarians. To allow doctors to access the developed tools, the VetWeb web portal was prepared. The portal enables remote access to artificial intelligence-based automatic diagnostics. The doctors can use the portal to submit the microscopic images of the cytological samples. When the analysis of the image samples made by the algorithm is ready, the doctor receives a report with information about pathological changes. The interactive report, available on the web portal, contains all the detailed information provided by the veterinarian during the sample data submission accompanied by photos with tumor sites marked on them. Also, the portal allows the doctors to track the history of the patients' treatment as data is stored in database.

Time series analysis

The works of the Laboratory also concern the modelling of time series to detect unusual situations. Work includes practical applications such as anomaly prediction to avoid catastrophic damage to the magnets and other associated devices of the LHC accelerator at CERN. Moreover, the developed algorithm was implemented on the Xilinx Zynq UltraScale+ MPSoC FPGA architecture to achieve a low latency response.

In the field of medicine, we address the problem of prediction of fainting of hospitalised patients, who are confined to a hospital bed for a long time. Thus, the analyses concern the well-known problem of losing the leg muscles supports for the cardiovascular system. This research is carried out with the cooperation of the Medical University of Graz.



Contact: Paweł Russek, p. russek [at] cyfronet.pl

Cyfronet Open Day 2022

Every year, the Academic Computer Center Cyfronet AGH organises an Open Day for the scientific community of Krakow and the Małopolska region.

The program of the event includes:

- presentation of Cyfronet's activities and information about hardware resources, software and services offered,
- presentation of the latest trends in the development of computer and network architectures, including resources that allow performing quantum calculations,
- speeches on the development and use of artificial intelligence methods and the computational capabilities of supercomputers and their application for the analysis of research results, simulations, visualisation and parallel analysis of huge data sets (big data),
- announcement of the results of the competition for the best doctoral thesis based on the Cyfronet computing resources and presentations of the winners.

Participants of the Open Day have the opportunity to get acquainted directly with the unique services of Cyfronet and the possibility of using them in research for the synergy of science and economy.


ATHENA
moc dla polskiej nauki
CYFRONET AGH
EPISODES PLATFORM
Platforma


AULA
ATHENA
REJESTRACJA
ZAPISY NA ZWIEDZANIE SERWEROWNI
PRACUJ W CYFRONECIE

HPC Users' Conference (KU KDM) 2023

The basic premise of the HPC Users' Conference was to initiate annual scientific meetings devoted to users performing computations in ACC Cyfronet AGH with use of high-performance computers, computing clusters and installed software.

The conference was launched in 2008 and included several presentations by Cyfronet employees – describing the resources available in the Centre, as well as numerous lectures of researchers presenting the scientific results achieved using Cyfronet hardware and software. In addition, two invited speeches were given – by Norbert Attig from Jülich Supercomputing Centre and Jaap A. Kaandorp from University of Amsterdam.

The first edition of the conference attracted much attention and increased the interest of users in Cyfronet resources. It proved that this type of event was much awaited and needed.

Nowadays, the HPC Users' Conference focuses on the large-scale computations and simulations, novel algorithms in computer science, tools and techniques relevant to high-performance computing, teaching in computer science, databases. However, the main aim of the conference is the overview of research results carried out using the computer resources of Cyfronet. It is also an opportunity to familiarize the users with the Centre and its resources, including the PLGrid infrastructure.

The conference includes a series of talks by scientists who perform research using Cyfronet resources and can present the role of these resources, typical usage scenarios and performance aspects. The event is an important opportunity for Cyfronet representatives to meet with these scientists and acquire the knowledge necessary to take the proper actions in order to adapt the computing infrastructure to scientists' needs and fulfil their requirements. On the other hand, the conference also gives a possibility for researchers representing various disciplines to exchange experience and become familiar with the new technologies and services currently being deployed at the Centre.

The crucial parts of the conference are meetings with suppliers of Cyfronet hardware and software, as well as the panel discussion on efficient use of these resources. The latter is always attended by users – researchers, who use the chance to get familiar with news regarding the computing infrastructure in the Centre and to inform Cyfronet experts about issues encountered while interacting with this infrastructure.

Contributed papers elaborated on the basis of the best conference talks are published in one of two well-regarded IT journals: Computing and Informatics (CAI) (http://www.cai.sk) or Computer Science (CSCI) (www.csci.agh.edu.pl).

https://www.cyfronet.pl/kukdm23

Celebration of the 50th anniversary of Cyfronet

On March 24, 2023, in the AGH auditorium in the A0 building, a meeting was held to celebrate the Jubilee of the 50th anniversary of Cyfronet.

The event was attended by the President of the Republic of Poland Andrzej Duda, Deputy Head of the Chancellery of the President of the Republic of Poland Piotr Ćwik, Voivode of Małopolska Łukasz Kmita, Vice-Marshal of the Małopolska Voivodeship Łukasz Smółka, Deputy Mayor of the City of Krakow Bogusław Kośmider, Rector of AGH University of Krakow prof. Jerzy Lis, Vice-Rector of the Jagiellonian University prof. Armen Edigarian, rectors and vice-rectors of universities, director of PSNC Cezary Mazurek PhD, Vice-Chairman of the PIONIER Consortium Council Maciej Stroiński PhD, members of the PIONIER Consortium Council and science, research and innovative economy partner units, as well as current and retired employees of ACC Cyfronet AGH.

During the ceremony, the President of the Republic of Poland gave orders and state decorations to Cyfronet employees.

During his speech, prof. Kazimierz Wiatr, Director of Cyfronet, presented the most important facts from the history of the Centre: from its establishment in 1973, through a series of events and dynamic technological changes of the last half-century, to the present state and further challenges. ICT resources of ACC Cyfronet AGH were also presented, including ICT network, supercomputers and data storage systems. After taking a commemorative photo, the anniversary cake was cut. The meeting ended with a music concert and the presentation of the Album of the 50th anniversary of Cyfronet.

The speech of President Andrzej Duda during the Jubilee of the 50th anniversary of Cyfronet

Decorated employees of Cyfronet

President Andrzej Duda receives from the Cyfronet Management the commemorative medal and Intel Itanium2 processor from the Baribal supercomputer

Souvenir photo on the cloisters

The unveiling of a commemorative plaque in the Cyfronet building at Nawojki 11

LUMI Day Poland 2023 conference

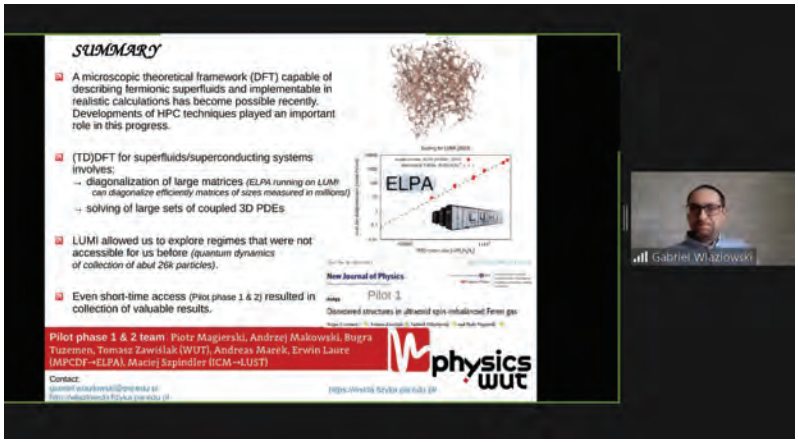
On April 14, 2023, an online conference was organised to present the LUMI supercomputer, the path of applying for its resources and the possibilities of its use by Polish scientists.

The LUMI supercomputer, built by a consortium covering 10 countries, including Poland, has been ranked 3rd on the TOP500 list of supercomputers with the highest computing power in the world.

Details about LUMI can be found at: *https://cyfronet.pl/lumi*.

Polish scientists can access LUMI via the national PLGrid infrastructure (*https://portal.plgrid.pl/*).

Topics covered at the conference included:



Prof. Gabriel Wlazłowski, the manager of the Polish project implemented using LUMI

- LUMI's place in the European EuroHPC JU initiative,
- scientific challenges to be answered thanks to LUMI,
- application path for LUMI supercomputer resources,
- support for LUMI supercomputer users,
- software in a pre-exascale environment,
- examples of using LUMI supercomputer resources,
- EuroCC – National Competence Centers in the framework of EuroHPC.

LUMI supercomputer, photo: Fade Creative

ACC Cyfronet AGH at the ISC High Performance 2023 in Hamburg

Cyfronet acted as an exhibitor during one of the most important events for the environment of suppliers and users of high-scale computing solutions.

ISC High Performance (May 21-29, 2023) attracted many exhibitors and about 3,000 participants from all over the world. The exhibitors included companies – technology giants, such as Hewlett Packard Enterprise, Intel or NVIDIA, as well as leading initiatives and supercomputing centers, including i.e. EuroHPC JU, LUMI, Barcelona Supercomputing Center. The stand of ACC Cyfronet AGH was one of the few Polish accents during the fair. The main activities of Cyfronet included: maintaining relationships with existing partners, establishing new contacts with potential hardware and software suppliers, and assessing new technological solutions that can be implemented as part of the managed infrastructure.

Cyfronet's participation in ISC was also related to promotional activities under the projects: "National Competence Centers in the framework of EuroHPC Phase 2" and "National Supercomputing Infrastructure for EuroHPC – EuroHPC PL".

Cyfronet AGH honored with the Polonia Minor Award

On May 31, 2023, during the meeting of the Małopolska Council for Information Society (MRSI), Cyfronet AGH received the Polonia Minor Award.

The Polonia Minor Award is one of the most important distinctions awarded by the Marshal of the Małopolska Region and the Chairman of the Małopolska Regional Council. The medals have been awarded since 2011 for outstanding achievements in the field of disseminating and popularising the idea of self-governance and for activities for the benefit of Małopolska region. Polonia Minor has the form of a medal, the obverse of which depicts Prince Bolesław the Chaste and his wife St. Kinga and St. Stanislaus BM, while the reverse shows the coat of arms of the Małopolska region inscribed in the contour of its contemporary borders. The author of the bronze cast is a Krakow sculptor, prof. Wincenty Kućma.

The Polonia Minor award was presented to the Director of Cyfronet AGH, prof. Kazimierz Wiatr by the Deputy Marshal of the Małopolska Region, Józef Gawron, on the occasion of the 50th anniversary of Cyfronet AGH, „in gratitude for professional activity for the development of IT infrastructure and many years of service and work for Polish and Małopolska science".

– *It is impossible not to mention on this occasion the extremely important role played by the Academic Computer Centre Cyfronet AGH in the development of Polish and Małopolska science, as well as innovative economy. On behalf of the region's authorities and the chairman of the Małopolska Regional Council, I would like to thank you for this invaluable contribution. At the same time, I wish you further success in attracting scientific talents and carrying out outstanding research works* – said Deputy Marshal Józef Gawron.

The diploma accompanying the medal was signed by the Marshal of the Małopolska Region, Witold Kozłowski, and the Chairman of the Regional Council of the Małopolska Region, Jan Duda.

The Małopolska Council for Information Society was established in 2004 and since then it has been an advisory and consultative body to the Małopolska Region Board in the field of:

- stimulating and monitoring the development of the information society in the Małopolska region,

- integrating projects for the development of the information society,
- shaping cooperation between public administration units in the region in the field of modern, digital administrative solutions,
- monitoring activities in the field of public information in the region.

The meeting of the MRSI was organised at the headquarters of ACC Cyfronet AGH. After adopting relevant resolutions by the Council, prof. Wiatr presented the most important facts from the history of the Centre since its establishment in 1973, as well as current ICT resources. After Cyfronet AGH was honored with the Polonia Minor award, the gathered representatives of the Council visited the Nawojki Data Center, getting familiar with the installed supercomputers, data storage systems and accompanying infrastructure.

At the end of the meeting, the guests signed in the commemorative book of the 50th anniversary of Cyfronet AGH.

Sat4Envi – space technologies, widespread use

By providing satellite data from the Copernicus program, the Sat4Envi project has opened a new chapter for optimal decision-making in spatial planning and rapid response to environmental threats.

Satellite data at your fingertips

The core achievement of the project is available at *dane.sat4envi.imgw.pl*. The Customer Service System is a web browser operated tool for accessing the Sat4Envi satellite data warehouse. Thanks to this platform, the end-user gains the ability to search, view, order and download satellite data and its derivatives. The system is primarily used to monitor the earth's surface, the atmosphere and the weather, and identify meteorological threats. Access to all data is completely free of charge and does not depend on affiliation or profession. Moreover, it does not require registration in the basic version. After logging in, however, the user has the option of using additional functionalities, such as saving a set of the most frequently selected measurements or accessing the metadata with which the presented satellite images are described.

By the possibility to use WMS (Web Map Service) layers, which include borders of administrative units, the tool is functional for public administration units or regional institutions. The operation of the Customer Service System is intuitive and allows to quickly reach the desired types of data.

The project and collaboration

A system for collecting, sharing and promoting digital satellite information about the environment – Sat4Envi – aimed at providing wide and easy access to satellite data from the Copernicus program and data from other environmental or meteorological satellites. Behind the success of Sat4Envi stands the involvement of four closely cooperating units. Those are: Institute of Meteorology and Water Management – National Research Institute (IMGW, the project leader), Polish Academy of Science Space Research Centre (CBK), Polish Space Agency (PAK) and Academic Computer Centre CYFRONET AGH-UST.

IMGW was primarily responsible for the development of the station for receiving and processing data from polar satellites (the station was established in Krakow) and played the role of the coordinator. The joint effort gave the possibility to efficiently implement the other assumptions of the project, including:

- building a modern satellite data archive (IMGW, CYFRONET),
- the training center for new satellite technologies (CBK, PAK, IMGW),
- the center for providing scientific information to users (CYFRONET, IMGW, CBK).

The project, with a total value of PLN 17,903,900, was co-financed by the European Regional Development Fund under the Digital Poland Operational Program.

Technological challenges

Building a portal with an access interface required a very careful and multi-faceted design, taking into account the ease of data search or the speed of finding specific functions. With the goal of creating a highly functional product also for "non-advanced" users, the development process was started with an extensive requirements assessment. Use cases were analyzed in various target user groups and multiple verifications were carried out using graphical mockups and prototype implementations. When working on the interface, user experience specialists also used the technique of in-depth interviews in groups of representatives of future users (i.e. IMGW, CBK, PAK, The Remote Sensing Laboratory or Pieniny National Park).

- *The role of Cyfronet in the project was to implement programming tasks related to the construction and maintenance of the Customer Service System, but also to provide the IT infrastructure necessary to collect and share data. Cyfronet is the operator of the Internet connection with the Institute of Meteorology and Water Management in Kraków, and as part of the PLGrid infrastructure, it provides computing resources necessary for the proper functioning of Sat4Envi. In turn, the Prometheus supercomputer installed in Cyfronet is used to process data and create information from raw satellite data* – emphasizes Prof. Kazimierz Wiatr, Director of ACC Cyfronet AGH-UST.

Sat4Envi supports in crisis situations

The analysis of satellite data with the smallest possible delays in relation to its acquisition may be of key importance for forecasting, monitoring and counteracting the effects of natural disasters. In particular, we are talking about monitoring floods, assessing the scale of fires, or monitoring damage to agricultural crops caused by drought or frost. Due to the importance of these activities and the need for rapid response, the Sat4Envi project includes support for special user groups such as the Crisis Information Center.

More information about the project can be found at: *https://sat4envi.imgw.pl/*.

Prometheus and its infrastructure are used to share data from Copernicus Programme. Image: Sat4Envi

EPOS – multidisciplinary platform supporting Earth sciences research

We need to use various tools and research methodologies to effectively monitor the processes ongoing below the Earth's surface and their effects perceptible on the surface. In this regard, EPOS – European Plate Observing System is to deliver complex solutions, which would cross the borders of different countries and research disciplines as well.

In October 2018, EPOS was granted the status of ERIC: European Research Infrastructure Consortium. This way legal and administrative framework was established, which allows EPOS to carry out international activities not only as a project implemented by many partners, but rather as a separate entity with its headquarters in Rome.

The idea standing behind EPOS is to recognise better Earth functioning as a complex system in which, on the one hand, natural episodes (such as volcanic eruptions, floods or earthquakes) have an impact on society and the economy. On the other hand, the environment is being changed by anthropogenic factors. For this reason, effective research requires a multifaceted approach. By integrating European research infrastructures into the transnational system, which delivers data, data products, software and services, EPOS works on enabling access to so far dispersed possibilities. The initiative focuses on ten main, linked with each other, Thematic Core Services (TCS): Seismology, Near-Fault Observatories, GNSS Data and Products, Volcano Observations, Satellite Data, Geomagnetic Observations, Anthropogenic Hazards, Multi-Scale Laboratories, Geo-Energy Test Beds for Low Carbon Energy, Tsunami.

Image: EPOS - European Plate Observing System

Computational Earth sciences

Providing the research infrastructure as an internet platform requires close cooperation of IT specialists with representatives of Earth sciences. Synergic activities are carried out on many levels – starting from measurements and experiments conducted by scientists generating a large amount of raw data, which is stored and processed in information systems – i.a., within Cyfronet's infrastructure. Cooperation is also significantly influential in building data visualisation tools, which are managed from the web browser level using Graphical User Interfaces (GUI). In the long run, EPOS aims to connect research communities within one multidisciplinary platform, provide effective and safe access tools, and develop new services based on previous achievements.

It's worth underlining that geophysical and geological data, advanced visualisation and analytic software will be available in open access and free of charge. Thanks to that, EPOS supports interdisciplinary research on the causes and effects of processes ongoing below and on the Earth's surface.

This is especially important in the light of monitoring environmental threats, such as earthquakes, floods, landslides or volcanic eruptions, and anthropogenic threats related to, among other things, the activities of mines. Through a comprehensive analysis of already observed phenomena, one can better prepare for events that have not yet taken place, limiting their adverse effects as much as possible.

EPOS has entered the operational phase of the infrastructure built within EPOS PP and EPOS-IP projects. The EPOS research platform is available at: *https://www.ics-c.epos-eu.org/,* and the portal gathering information about EPOS activities at: *https://www.epos-eu.org/*.

Cyfronet's work for EPOS

ACC Cyfronet AGH participates in subsequent EPOS projects, supporting the initiative with hardware resources and the specialists' knowledge and skills. The parallel activities include several areas, one of which is the development of the EPISODES digital platform for the Thematic Core Service – Anthropogenic Hazards (TCS AH). The platform, accessible at *https://episodes.eu,* is a tool for the analysis of anthropogenic seismicity and associated threats and for the assessment of the potential environmental impact of the exploitation of geo-resources. Thanks to data from seismic stations and shared industrial information, the platform allows more straightforward analysis of such processes as flooding artificial reservoirs, extraction of raw materials, shale gases or groundwater.

In addition to activities related to sharing and visualisation of data, Cyfronet's specialists take care of the maintenance of integration elements between the TCS AH and the superior EPOS ISC-C Portal, which gives access to the results of activities led in all of the EPOS Thematic Core Services.

Additionally, Cyfronet is sharing a virtual workspace for data management and for running software from various deliverers.

The expansion of the possibilities of conducting scientific research using the EPISODES platform is carried out as part of another project from the EPOS family: GEO-Inquire (Geosphere INfrastructures for QUestions into Integrated REsearch), which includes increasing the amount of data available on the platform, adapting the platform's capabilities for the needs of industrial research and integration platform services with European Open Science Cloud (EOSC) services.

Additionally, by participating in national projects, Cyfronet works to develop the Polish Research Infrastructure EPOS-PL. The latest effort, within the EPOS PL+ project, regards developing software that supports modelling and conducting research using AI methods. The EPOS-AI platform will enable easy use of data integrated into the program, selection of machine learning methods, and launching resource-intensive stages of processing on computational resources of Polish High-Performance Computing Centres.

Sano: Centre for Computational Personalised Medicine - International Research Foundation

Owing to a unique initiative carried out in 2019-2026 by the Academic Computing Centre Cyfronet AGH along with five partner institutions in the framework of the EU Horizon 2020 *Teaming for Excellence* programme, the International Research Agendas programme implemented by the Foundation for Polish Science, and with financial support from the Ministry of Science and Higher Education, a new entity called **Sano –** Centre for Computational Personalised Medicine was established in Kraków.

The mission of Sano involves:

- development of new computational methods, algorithms, models and technologies for personalized medicine,
- introducing new diagnostic and therapeutic solutions based on computerized simulations into clinical practice,
- fostering creation and growth of enterprises which develop cutting-edge diagnostic and therapeutic technologies,
- contributing to novel training and education curricula which meet the needs of modern personalised medicine.

The **Sano Centre** (*https://sano.science/*) is situated in Kraków: a city well known for educating top-class medical practitioners and IT experts, whose teaching hospitals are well regarded among the academic community and whose life science technology sector is continually expanding.

The establishment of the **Sano Centre** directly contributes to regional scientific excellence by fostering new research collaborations and creating top-tier educational opportunities for postgraduate students. It will also improve knowledge and technology transfer by promoting creation of new commercial enterprises which deal with advanced technologies. The Centre's impact will transcend regional boundaries, contributing to advancements in medical research and thereby to the quality of medical care.

The Centre's objectives are based, among others, on the National Smart Specialisation Strategy. **Sano** aims to enhance collaboration between academic and commercial institutions on an international scale. Key performance indicators include the number of highly cited scientific publications and grants obtained by the Centre, the number of solutions based on computational models which have been introduced into clinical practice, and the number of innovative marketable products and services.

The Centre for Computational Personalised Medicine represents a joint international collaboration of the following institutions: ACC Cyfronet AGH, LifeScience Cluster Krakow – a Key National Cluster, University of Sheffield and Insigneo Institute, Forschungszentrum Jülich, Fraunhofer Institute for Systems and Innovation Research ISI, and National Center for Research and Development.

Digital Twins

The concept of Digital Twins (DT) assumes the creation of a virtual model that reflects the features of the object adopted as a model as faithfully as possible. This makes it possible to conduct computer analyses to take effective, model-tested actions in reality. Through its infrastructure and expertise, Cyfronet supports creating the "Digital Twin" model in various fields of science.

The EDITH project (An Ecosystem for DIgital Twins in Healthcare) is developing an integrated system architecture to implement the human "Digital Twin" concept and a vision for its further development. The Cyfronet team is coordinating work on developing a demonstration prototype of the simulation platform.

In turn, the DT-GEO project (A Digital Twin for GEOphysical extremes) assumes the implementation of a "Digital Twin" prototype to study extreme geophysical conditions, such as earthquakes (natural or anthropogenic), landslides, volcanic eruptions and tsunamis. In this case, Cyfronet specialists are working on the part related to anthropogenic threats.

The InterTwin project (An interdisciplinary Digital Twin Engine for Science) co-designs and implements a prototype of an interdisciplinary solution called Digital Twin Engine (DTE). This open-source platform provides generic and customised modelling and simulation software components to integrate Digital Twins (DTs) specific to the application. In the project, the Cyfronet team introduces integration with the proprietary OneData platform.

The cloud environment

The ability to run a complete work environment in the cloud, with access to distributed hardware resources and software packages, facilitates research in international teams. Cyfronet has been developing cloud "ecosystems" for years based on available open-source tools and proprietary technologies.

The main goal of the INDIGO-DataCloud project (INtegrating Distributed data Infrastructures for Global ExplOitation) was to develop a PaaS (Platform as a Service) environment enabling large-scale computing by integrating grid and cloud resources and providing uniform access to both computing resources and data storage resources. The role of ACC Cyfronet AGH was mainly to implement a consistent system of access to geographically dispersed data stored in heterogeneous repositories.

Based on the project's achievements mentioned above, another project was carried out: XDC (eXtreme DataCloud), which aimed to build specialised solutions for large-scale data management and processing in a hybrid cloud, thus introducing the issues of data access and migration in distributed cloud environments. Cyfronet was responsible for implementing a distributed data management system in a cloud environment.

One of currently implemented projects is DOME (A Distributed Open Marketplace for Europe Cloud and Edge Services). Based on the assumptions of the Gaia-X program and open standards, tools and procedures are being developed to support the development and adoption of trusted Cloud and Edge services in Europe. The DOME project will be an access point for broadly understood software and data processing services developed under EU programs such as Digital Europe, Horizon 2020 or Horizon Europe.

EOSC – transnational integration of scientific resources

The European Open Science Cloud (EOSC) is an initiative aiming to create a virtual environment which would complete the assumptions of the Open Science paradigm. EOSC aims to share (easily and transparently) research data and advanced tools and resources to store, share, process and manage this data.

Within EOSC, the connections between currently existing research e-infrastructures are made, and the integration takes place, i.a., by the unification of access and authorisation rules for researchers from different countries. Thanks to the achievements of EOSC-Hub, belonging to the family of EU-funded EOSC-building projects, a platform was created to do this. The EOSC-Portal is the interface between providers of scientific services and resources, and researchers who can benefit from those. ACC Cyfronet AGH team played a crucial role in this regard, becoming the Portal host and developing the Marketplace website, an extensive catalogue of services and documentation provided by EOSC partners.

Further development of EOSC

Development of the Portal and integration of services from the larger group of deliverers are the core assumptions of the planned enhancement of EOSC.

The EOSC Enhance project (*https://www.eosc-portal.eu/enhance*), which was realised from 2019 to 2021, aimed to improve the Portal in terms of convenience and speed of use. The works carried out by our specialists concerned, among others, advanced analysis of user behaviour in order to create and implement the best user experience practices. At the same time, new functionalities have been implemented.

The EOSC Synergy project (*https://www.eosc-synergy.eu/*), which lasted until the end of October 2022, has been underway to implement EOSC standards for another nine national e-infrastructures. Cyfronet, in addition to coordinating activities at the national level, supported the planning process by looking for new, effective solutions for integration in other countries.

In turn, the EOSC Future project is a continuation and an enhancement of the work that has been done so far. The project aims to upgrade the EOSC ecosystem so that it supports European research even better, and to convince researchers to use the resources which are offered. Cyfronet's responsibility as part of EOSC Future is the further development of one of the three main pillars of the Portal, which is the part aimed at the users, in particular related to the components supporting the Marketplace.

Cyfronet also participates in the latest initiatives building EOSC: FAIRCORE4EOSC and EuroScienceGateway. In the first project, our role is to provide expert support in the field of data management and personalised search of EOSC resources (including scientific objects) using AI methods. As part of EuroScienceGateway, Cyfronet is responsible for integrating the solutions of the Onedata distributed data management platform with the project infrastructure being built.

EuroHPC – European Data-Processing Infrastructure

The European High-Performance Computing Joint Undertaking (EuroHPC JU) was established to radically develop the existing European HPC infrastructure so that it could provide European researchers with computing power comparable to the ones available in the USA, China and Japan. EuroHPC unites 33 countries, from 2018 including Poland, as well as private members. The joint aim is to buy and deploy two exascale supercomputers (EFlops = 10^{18} floating points per second) that will be on the TOP5 list of the world's fastest computers.

The first European exascale supercomputer will be JUPITER at the Jülich Supercomputing Center in Germany, which will become the fastest European machine when fully installed in 2024. The second supercomputer with a power exceeding 1 EFlops will be located in the TGCC computer centre in France. Before that, however, subsequent lower-power supercomputers are launched, creating an infrastructure that allows the technology and software to be scaled. By the time this publication was prepared, EuroHPC JU had delivered 3 pre-exascale supercomputers: LUMI (Finland), Leonardo (Italy), Mare Nostrum 5 (Spain). Additionally, 5 petascale supercomputers were successfully put into operation: Discoverer (Bulgaria), Karolina (Czech Republic), MeluXina (Luxembourg), Vega (Slovenia) and Deucalion (Portugal). The next in this class of computers will be Daedalus (Greece). EuroHPC JU also finances the purchase of quantum computers. Polish scientists will have access to one of them, installed in the Czech Republic, via the PLGrid infrastructure coordinated by Cyfronet.

More about EuroHPC JU at: *https://eurohpc-ju.europa.eu/*.

Karolina Supercomputer, Czech Republic. Image: EuroHPC JU

Vega Supercomputer, Slovenia. Image: Atos

EuroHPC PL – the Polish part of the EuroHPC JU

At the beginning of 2021, the National Supercomputing Infrastructure for EuroHPC JU – EuroHPC PL project was launched. It aims to deliver the modern infrastructure, particularly supercomputing resources, specialised accelerators and laboratories' services for academia, industry and society.

The project is run by the Consortium consisting of 7 members:

1. AGH University of Science and Technology in Krakow - Academic Computer Centre CYFRONET – Consortium Leader,
2. Institute of Bioorganic Chemistry of the Polish Academy of Sciences - Poznań Supercomputing and Networking Center,
3. Gdańsk University of Technology - IT Center of the Tricity Academic Computer Network,
4. Wrocław University of Science and Technology - Wrocław Center for Networking and Supercomputing,
5. National Center for Nuclear Research,
6. Institute of Theoretical and Applied Informatics of the Polish Academy of Sciences,
7. Center for Theoretical Physics of the Polish Academy of Sciences.

EuroHPC PL was established to facilitate and expand access to large-scale data processing for both scientists and entrepreneurs and thus increase the competitiveness of Polish entities compared to Europe and the world. The essential element of the infrastructure being built is the computing power provided by supercomputers. Their use is a crucial tool for conducting scientific research in many fields and significantly accelerates and reduces the costs of research processes. It is especially visible at the initial stage of the research process, in which the set of possible research directions is so large that their classic verification (e.g. in laboratories) would not be possible due to both time consumption and costs. Supercomputing resources will be supported by the architecture of quantum algorithms and neuromorphic accelerators. This will significantly increase the possibilities of data analysis and visualisation. Additionally, the project provides a wide range of user support and training.

Project tasks are performed within four laboratories. The main goal of the Laboratory of Modeling and Parallel Data Processing in a Pre-Exascale Environment is to design and implement advanced large-scale computing systems. The Laboratory for the Application of Hybrid Computing addresses the needs related to the application and implementation of practical use of quantum computers. In turn, the tasks of the Laboratory for the Application of Supercomputers in Medicine focus on the practical use of high computing power in medical applications. The Laboratory of Energy and Computing Efficiency of HPC Software works on the development of techniques to reduce machine operating costs, such as energy consumption and cooling demand, as well as works on optimising software running in a supercomputing environment.

More information about EuroHPC PL is at: *https://eurohpc.pl*.

The LUMI supercomputer

LUMI - No. 1 in Europe, No. 3 in the world

In June 2023, during the ISC High Performance fair in Hamburg, another TOP500 list of supercomputers with the highest computing power in the world was announced.

The LUMI supercomputer installed in the CSC data centre in Finland took third place there while maintaining the lead in Europe. This undoubted success in the HPC (High-Performance Computing) environment is the result of the cooperation of 10 countries, including Poland, which formed a consortium operating under the wings of a larger initiative, the EuroHPC Joint Undertaking. Importantly, thanks to the activities coordinated by ACC Cyfronet AGH and the Ministry of Education and Science, Poles have already gained access to LUMI resources.

LUMI among the world leaders in computing power and energy efficiency

Together with other petascale and pre-exascale systems in which EuroHPC JU is involved, LUMI is a part of the pan-European infrastructure to which fully exascale computers will also be incorporated in the coming years. Obtaining a place at the TOP3 of supercomputers is a great success for European countries because it shows the possibility of competing with systems from the USA, China and Japan. In High-Performance Linpack (HPL) tests, the LUMI supercomputer showed an actual power of 375 PFlops and a theoretical performance of up to 550 PFlops. This performance is made possible thanks to the LUMI hardware architecture:

- the GPU partition is composed of 2560 nodes, each of which contains a 64-core AMD Trento processor and four AMD MI250X cards,
- a single MI250X card can deliver 42.2 TFlop/s of performance in High-Performance Linpack tests,
- each GPU node contains four 200 Gb/s network connection cards,
- the addition to the GPU partition is the CPU partition, which uses the 3rd generation AMD EPYC™ 64-core general purpose processors.

However, during the design and construction of the LUMI infrastructure, a strong emphasis was put not only on the computing power, but also on the efficient use of energy. For this reason, the LUMI supercomputer also took third place on the Green500 list of supercomputers with the highest energy efficiency, calculated as the ratio of computing power to electricity consumed. In addition, it is worth emphasizing that 100% of the LUMI's power supply comes from renewable sources (hydroelectric power plants), and the waste heat removed from the machine is used to heat the surrounding buildings. Thus, LUMI is now a flagship European example of implementing the idea of sustainable development in the construction of supercomputing systems.

More information about LUMI at: *https://lumi-supercomputer.eu*.

The representatives of ACC Cyfronet AGH during the inauguration of LUMI.
From the left: Mariusz Sterzel, Head of the Quantum Computing Laboratory and Marek Magryś, Deputy Director of ACC Cyfronet AGH for HPC

Access to LUMI for Polish scientists

ACC Cyfronet AGH coordinates work related to providing Polish scientists with the resources offered by LUMI. The aim is to ensure an efficient and user-friendly environment for working with a supercomputer, based on good practices developed within the PLGrid infrastructure.

In the period from autumn 2021 to spring 2022, Cyfronet organised two competitions for pilot computing grants on the LUMI supercomputer for Polish scientists. Subsequent competitions concern regular, annual access to infrastructure. Based on them, calculations with the use of LUMI have already been carried out by 15 Polish projects. Access is provided via the PLGrid Portal (*https://portal.plgrid.pl*), and additional information is available at *https://cyfronet.pl/lumi*.

Cyfronet is the leader of projects included on the Polish Research Infrastructure Map

Among the strategic infrastructures included in January 2020 on the Polish Research Infrastructure Map there are two projects proposed by ACC Cyfronet AGH as the initiator and coordinator of the PLGrid consortium: *National Supercomputing Infrastructure for EuroHPC* and *National Cloud Infrastructure PLGrid for EOSC*.*

The aim of the **National Supercomputing Infrastructure for EuroHPC** program is to build a computing infrastructure for scientific research on solutions that meet the current and future needs of Polish society, the scientific community and the economy.

The infrastructure will be based on modern supercomputing systems enabling the implementation of both traditional simulation tasks and data analysis using artificial intelligence methods. The production computing systems built within the project will be among the world's leading supercomputers. In addition to the computing infrastructure, the project will also provide access to specialized training and expert technical support for users from science and economy, as well as the necessary procedures concerning allocation and accounting of used resources.

The project fits directly into the framework of the international EuroHPC – European High-Performance Computing initiative published in the European Commission Communication COM/2018/08 final – 2018/03 (NLE). EuroHPC is a project aimed at creating a European system of high-performance exascale computers, unique on a global scale, based on technologies developed in Europe. The project will be implemented by the PLGrid Consortium.

As a result of the implementation of the National Supercomputing Infrastructure for EuroHPC project, an infrastructure for conducting research for the needs of science, economy and society will be created, benefiting from the latest HPC technologies developed within the international EuroHPC cooperation. The infrastructure will offer services in the fields of massively parallel computer simulations, highly efficient processing of data sets, the use of artificial intelligence methods, software and high productivity tools, including data visualization, and user support and training. In addition to the main computing systems, the infrastructure will also include smaller test and research systems to verify new processor, accelerator, memory and network technologies in the context of using them to build production systems, as well as conducting research and development in the field of effective HPC infrastructures.

The services will be offered via the PLGrid infrastructure, which integrates most of the computing resources available in the country, what will facilitate the process of resource allocation and user support. Integration will also be implemented with European pre-exascale and petascale systems made available under the Euro-HPC program, in particular with a machine built by the LUMI consortium, of which Poland is a member.

National Cloud Infrastructure PLGrid for EOSC is a program for the use of cloud resources for scientific research that meets the current and future needs of the Polish society, the scientific community and the economy. The scope of this research includes data, infrastructures and data processing platforms, as well as effective algorithms and dedicated applications.

The program is based on the requirements of the society, economy and Polish researchers, in particular those cooperating within international research groups. These groups require advanced environments for the integration of distributed resources: software, infrastructures and dedicated services. These requirements can only be met by advanced IT technologies combined with computing, storage and data resources. Cloud technologies enable the interaction of all these elements within a flexible ecosystem.

National Cloud Infrastructure PLGrid for EOSC is part of the ecosystem of the European Open Science Cloud (EOSC, Declaration of 26.10.2017). Poland is currently developing two key components of this federated, globally available and multidisciplinary environment: Onedata – a system for unified data sharing and management, and the EOSC Portal. As part of the European ecosystem, the PLGrid National Cloud Infrastructure will offer trusted and open environments for users throughout the data lifecycle. This will allow scientists, the economy and society as a whole to publish, search, use and re-use the collected data, tools, software and other results.

The research planned within the National Cloud Infrastructure PLGrid for EOSC will allow for the development, validation and, consequently, the provision of services (general and dedicated), and thus the use of modern technologies and effective techniques for management, processing and reusing data by scientific communities, economic entities and society.

To this end, the Infrastructure will provide:

- solution technologies for distributed environments, including cloud environments, covering service management in a distributed environment, automation of complex processes, integration of research platforms and infrastructures,
- safe data sharing and management based on national technologies, in accordance with EOSC standards,
- research on the specific needs of users in the field of large-scale data processing in a distributed environment, including: „close data" processing, using the „data lakes" paradigm with new analysis models, scalable resources in a distributed environment,
- verification of solutions prepared for the needs of society, science and economy in advanced applications,
- a catalog of general cloud services for the economy, science and society as a result of research and development works in cooperation with EOSC.

The first stage of this work is currently carried out within the resources of the PLGrid infrastructure.

**Material from the "Polish Research Infrastructure Map" brochure of the Ministry of Science and Higher Education.*

The main goal of the project is to provide an open and fully operational FAIR EOSC ecosystem. The project aims to expand the EOSC infrastructure with components supporting the FAIR paradigm (Findability, Accessibility, Interoperability, Reusability).

The aim of the project is to define and develop a new and innovative European Master's degree program focusing on HPC. The program is designed to equip students with HPC competences and knowledge required by academia and industry.

The PROTEUS-RS project aims to implement models and a number of numerical simulations based on high-performance computer architectures for the design of the manufacturing process of metal long elements.

Through the development of the EOSC Portal, the project aims to extend the EOSC ecosystem, integrate existing scientific infrastructures and initiatives, and engage new domain and national research communities.

The main aim of the project was to deliver transparent and unified access to Earth Observation Data, primarily satellite data, in order to support food production and food and agriculture industry.

The goal of the project is to create a computational medicine centre in Krakow. The Centre will be the main driver of European progress in this fast-growing sector, developing advanced engineering methods for the prevention, diagnosis and treatment of diseases, and meeting the global need for radically improved healthcare systems.

The aim of the project was to expand the computing infrastructure enabling cooperation in the area of high-intensity problems connecting large volumes of data, which cross the boundaries of a single data center.

The purpose of the EPOS-SP project was to develop and implement assumptions to ensure the sustainability of the EPOS infrastructure produced in previous projects (EPOS-PP and EPOS-IP).

The EOSC-Synergy project introduced EOSC standards for national infrastructures in nine European Union countries. It was done by harmonizing policies and expanding access to research infrastructures, scientific data and domain services.

The goal of the PRACE-6IP project was to implement new solutions and maintain the operationality of the PRACE environment in the area of European HPC computing infrastructures.

The PRIMAGE project aimed at creation of a Clinical Decision Support System (CDSS) for the treatment of cancer (neuroblastoma, glioma) in children. Patients' data was used in the multi-scale computational models of cancer designed to define disease biomarkers. The created CDSS system helps oncologists both in diagnosis and in predicting of disease progression and treatment effectiveness.

The mission of EuroCC 2 is to continue the creation of the EuroHPC National Competence Centers (NCCs) network in Europe through cooperation, exchange of best practices and knowledge and upgrading the national and European level of HPC service delivery.

The EOSC Enhance project aimed to build an improved, more integrated version of the EOSC Portal, which enables improvement and extension of solutions that make it easier to find European scientific services and open science data sets.

The aim of the EOSC-hub project was to prepare the launch of a production infrastructure for open science in Europe and the practical application of solutions developed as part of the EOSC-Pilot project to a real large-scale environment scattered across most European countries.

The aim of the project was to build a specialized general-purpose infrastructure for large-scale computing, enabling the undertaking of research challenges in key areas from the point of view of Polish society, the scientific community and the economy. The project was the Polish stage of development of the EuroHPC program.

The goal of the project was to develop and provide production services for storing, accessing, securing data and managing metadata, as well as integrating solutions for processing large and complex data volumes on the basis of a distributed e-infrastructure.

Within the project the construction of unique research laboratories based on the national PIONIER fiber optic network was realised. The main goal of the project was to build and make available platforms for research units, entrepreneurs and other entities interested in conducting scientific research and development works based on a new, nationwide research infrastructure.

The direct goal of the project was to create a specialized e-infrastructure for data processing, enabling the optimal use of specialized and new generation services to stimulate new areas of application in science, economy, education and social life.

As part of the project, the functionalities of the EPOS-PL research infrastructure were increased. A new Research Infrastructure Center (Center for Research Infrastructure of Satellite Data – CIBDS) was established, a new test site (Geophysical Safety System for mining protection pillars) was created, and the WNiP was established: „IT Platform for Research with Artificial Intelligence Methods (EPOS-AI)".

The goal of the project was to create computing infrastructure services and data storage services for the purposes of the PRACE project, within six dedicated laboratories: 1) L. of HPC and cloud processing, 2) L. of access to processing infrastructure, 3) L. of service management and monitoring, 4) L. of data management services, 5) L. of distributed data management and transparent access to data, 6) L. of infrastructure security.

The main objective of the project was to provide satellite data coming from the Sentinel satellites of the Copernicus network. The project created an infrastructure for automatically downloading data directly from satellites, its secure storage and sharing for the purposes of science, administration and training.

The project aimed at building the national research infrastructure for solid Earth Science and its integration with international databases and services implemented under the European Plate Observing System (EPOS).

The objective of the project was the development of the specialized technological competence centre in the field of distributed computing infrastructures, with particular emphasis on grid technologies, cloud computing and the infrastructures supporting calculations on large data sets. As a result, a great computing power and huge storage for digital data were offered to users. They also obtained access to a set of basic and end-user services.

The project aims to provide a quantum critical distribution infrastructure and a quantum communication network using the existing fibre infrastructure of the PIONIER network.

The project will build a cloud environment for content from the CH (Cultural Heritage) area, processing and collecting high-quality content, allowing the launching of virtual spaces for an interdisciplinary and cross-sector community cooperating with the CH area.

The project will provide new or significantly improved access to key observations, data products and services that enable the monitoring and simulation of dynamic processes in the geosphere at an unprecedented level of detail and precision, both spatially and temporally.

The laureates of previous editions of the Contest

Work of young scientists in Cyfronet

The annual contest for the best PhD thesis conducted with the help of computing resources of ACC Cyfronet AGH is a tradition in our Centre. The scientific value of the submitted doctoral dissertation is assessed, as well as the possibility of its practical application and the scope of use of computing resources and disk storage in Cyfronet. In recent years, the Contest has become an important event promoting research conducted by young scientists. To subsequent editions of the Contest, participants submitted many PhD theses focused on various scientific problems in chemistry, biophysics, physics and computer science. Also, the utilisation of the resources varies, as the contesters use different tools running on a wide range of computing architectures offered by Cyfronet.

The laureates of the Contest are invited to give a talk during Cyfronet's Open Day. We are honoured to present here selected interviews with the Contest participants.

Join the next Contest edition!
http://www.cyfronet.pl/konkurs

The laureates of the Contest in 2022

Karol Capała

The interview with the author of the PhD thesis:
"Anomalous Stochastic Dynamics in the Underdamped Regime"

What prompted you to become a scientist?

It's hard for me to put it into a story because I feel like I've always had this plan. I have always had a hunger for knowledge. When I decided in the middle school that I would like to become a physicist, I knew that it was inherently linked to staying in science.

So your interests developed very quickly.

Thanks to my dad, I have always enjoyed maths and physics. Their structured nature fits perfectly with my way of thinking. At the same time, I like the connection to the world that physics shows, so it was a natural choice for me.

My favorite subject during my studies became Classical Mechanics. It was taught by Prof. Bartłomiej Dybiec. I became interested in what Prof. Dybiec does scientifically, and it was he who made me fascinated with stochastic processes. And underdamped anomalous diffusion was a poorly researched topic at the time I started my PhD, so I decided to change this state.

Could you please outline the issues covered by your doctoral dissertation?

Stochastic processes constitute a vast field of research at the borderline between mathematics and physics. They can be dealt with in a purely mathematical context, for use in machine learning, or the creation of models of physical, chemical, biological, or socio-economic phenomena. My research is somewhere on the borderline between the first and the last category. These issues can be imagined using the motion of a particle. The particle can be any object so small that its shape is irrelevant to the problem under consideration, for example, the pollen of a plant. Let us imagine that our particle, our pollen, is in the air and can move freely in it. It is therefore described by its position and velocity. As we know, the air is made up of many particles, which are themselves in constant motion, constantly colliding and, more importantly for us, hitting our pollen, and thus able to change its velocity, both its value and its direction. Since there are more collisions in such a system than we would be able to describe with even the most powerful computers, it is easier to represent the system as our particle (pollen) and the noise, a random force that comes from averaging the collisions with the environment. Among other things, the study of such systems composed of particles and noise is dealt with by stochastic processes. Now the question of Lévy noise remains. The distribution of noise pulses, which are the average of collisions with the environment, is mostly Gaussian. However, this is not always true; the assumptions of the central limit theorem are not always fulfilled. This happens, for example, when, apart from collisions, the environment can also interact with the particle at a distance. The simplest example here would probably be the electromagnetic interaction – if our particle and its environment are electrically charged. In this case, the incident force with which the environment can

act on our particle is no longer limited and extreme events become quite probable, whereby the particle can suddenly experience a very large change in velocity. Lévy noise is an example of noise that can describe this type of behavior. Unfortunately, due to the ability to describe extreme phenomena Lévy noise is characterised with a lack of variance and, in many cases, also a lack of mean. Yes, we cannot even talk about the mean value of the noise in many cases. Dealing with these limitations is already an interesting mathematical problem, but I am afraid there is no place for such considerations here.

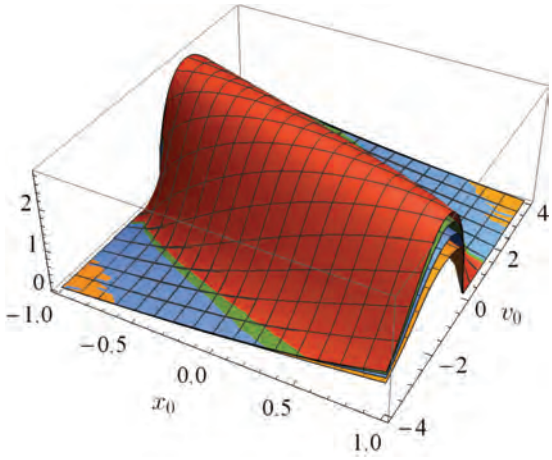
For my dissertation, it was important how these extreme events affect two classical problems in the field of stochastic processes: the steady-state problem, i.e. the probability distribution of finding a particle at a given point in space after a very long time, and the escape problem, i.e., as the name suggests, the study of how the process of exiting the studied area takes place for such a particle. The first of these issues is relevant to machine learning because of its connection to the stochastic gradient descent method, while escape problems are common in chemistry or biology. As part of my dissertation, I showed, among other things, that in a certain, broad class of potentials it is possible to have stationary states for which the most likely position is not the position with the lowest energy, and that the speed of transition between two states in a two-hole potential depends on both the distance between the states and the height of the energy barrier separating them.

You were performing computer calculations. How much of the work were they?

In my field, simulations, that can be almost completely parallelised, are an important part of the research. Therefore, high-powered computers are extremely useful, allowing one to conveniently obtain results for a wide range of parameters relatively quickly. I ran my first simulations on Cyfronet computers while I was still doing my undergraduate thesis. At that time I used the Zeus supercomputer, but while working on my PhD I already used Prometheus. In terms of the use of the software at Cyfronet, I was not particularly demanding, because the simulations I was running I prepared myself in C++. Of course, I would recommend working with supercomputers to anyone who needs to do a lot of calculations on a computer. This significantly shortens the waiting time for results.

What next?

I am planning my scientific development and habilitation is a natural next step in my career path. I am currently combining my research on stochastic processes with my work at Sano – Centre for Personalised Computational Medicine, where I am conducting research towards the creation of an artificial intelligence algorithm for medicine.



Mean first exit time from the (-1, 1) interval as a function of the initial position and velocity for various Lévy noises

You have recently completed a certain stage. Is there anything you would like to know today before starting your PhD?

I guess I'm lucky because the only thing I can think of is that I wish I had known that I was facing a lockdown caused by a pandemic after my first year of study – I could then have planned some elements of my project better. Any other knowledge that turned out to be useful or helpful was passed on to me either by my thesis' supervisor, Prof. Bartłomiej Dybiec or older colleagues. Therefore, I guess my main advice when choosing a PhD would be to be guided by both the topic and the atmosphere in the group.

Thank you.

Sylwia Gutowska

The interview with the author of the PhD thesis:
"Study of the electron structure, electron-phonon interaction and superconductivity of materials containing heavy elements"

I would like to start by asking about the source of your fascination with theoretical physics.

Like for everyone else, it's a desire for a deeper understanding of the surrounding reality. Theoretical physics provides the opportunity not only to recognise that certain phenomena occur, but also why and how they happen.

Since childhood, I have been more interested in mathematics, which I have treated as logical puzzles. My physics teacher in high school convinced me that physics is mathematics in practice. Additionally, I was interested in philosophy at that time, so studying the nature of the universe was intriguing to me.

When I started studying technical physics, I became interested in solid-state physics. I was fascinated by the fact that macroscopic properties of materials, such as conductivity, can be predicted from quantum mechanical equations like the Schrödinger equations. That's why I focused on the band theory.

What made you decide to devote your doctorate to the phenomena associated with superconductivity?

I have to admit that I chose my advisor more than the topic itself. My advisor was conducting classes in solid-state physics, and at that time I became interested in this field. Therefore, I decided to ask him to be the supervisor of my engineering thesis, and he suggested that I work on superconductivity. During my doctoral studies, I had the opportunity to continue exploration of this topic, already equipped with the necessary theoretical knowledge and familiarity with computational codes.

The topic itself is very intriguing to me. Despite the fact that the phenomenon of superconductivity has been known for over 100 years, there is still no universal recipe for a good superconductor. Thus, there are still many aspects to investigate and hypotheses to formulate.

Do you expect that one day we will see superconducting materials at room temperature?

In recent years, superconductivity under very high pressure has been discovered in systems containing hydrogen. The critical temperatures in these compounds already reach 250 K (-23°C), and it is worth emphasising that theoretical calculations have played an important role in these discoveries, confirming their validity and suggesting which systems should be experimentally studied. However, such materials are far from practical applications, as the technological challenge shifts from cooling to low temperatures to maintaining extreme pressures.

In March 2023 a report was published on the discovery of room-temperature superconductivity in doped lutetium hydride under relatively low pressure of 1 GPa (Dasenbrock-Gammon, et al., *Nature* **615** (2023)). However, this work has faced significant criticism from the scientific community. According to the scientific methodology, an experiment must be repeatable to be considered true. Until other groups are able to reproduce it, this discovery should be treated with caution.

I don't know if achieving room-temperature superconductivity at atmospheric pressure will be possible, but I believe temperatures above 200 K are attainable (the current record under atmospheric pressure is 138 K).

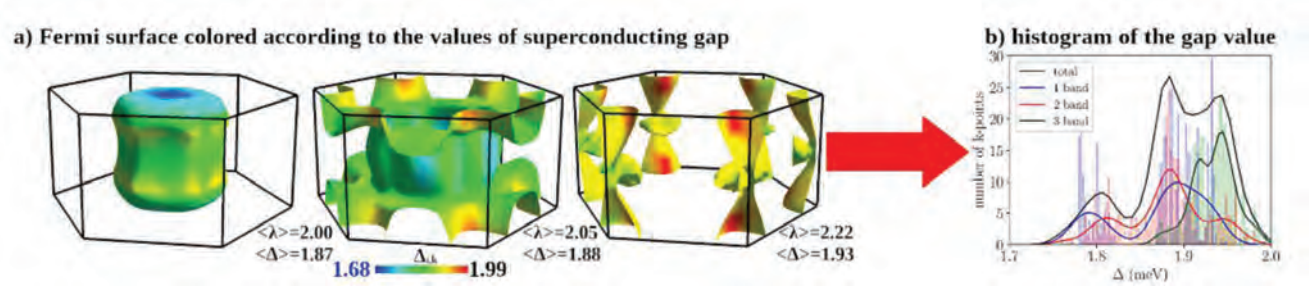
Did Cyfronet provide you with all the necessary tools to perform the calculations?

Yes, mostly. I conducted the most complex calculations on Zeus and Prometheus supercomputers. The ability to simultaneously perform calculations on hundreds of processors and the access to pre-installed software provided significant support.

You have an impressive number of publications in worldwide journals – how do you assess the role of scientific discourse at this level? Any tips for people who would also like to publish?

I think that collaboration with other people is crucial. In our case, we closely collaborate with research groups conducting experimental investigations, which greatly supports our computations. This provides an opportunity for broader discussions of the results, assures us that our theoretical analysis is accurate, and explains the experimentally observed phenomena. Ultimately, this makes our publications more interesting for readers and facilitates their acceptance by editors and reviewers.

Thank you for interview.



The anisotropic distribution of the superconducting gap values in the Pb-Bi alloy with exceptionally strong electron-phonon interaction, which is the cause of superconductivity. It has an impact on the thermodynamic properties of the superconducting state observed in experiments with this compound, including magnetic critical field, critical temperature, and specific heat

Tomasz Kura

The interview with the author of the PhD thesis:
"Numerical analysis of heat and mass transfer in the zone of mini-jets interaction with cylindrical surface"

What made you choose your academic path and specific specialisation?

My studies began coinciding with the period when the energy transformation was talked about almost everywhere. That's why I chose Energy as a good and potentially future-oriented field of study. However, I was surprised to discover that there is such a field of science as numerical methods and simulations. When I had the opportunity to learn more during my studies, this knowledge from the borderline of the world of IT and classic energy seemed worthy of attention. After a few years, I can confirm that it was a good choice.

How would you explain to a non-technical person how a mini-jet heat exchanger works?

The exchanger has unique elements, tubes with over a thousand tiny holes. Their location close to the surface that separates the heat receiving and releasing factors means that these factors (e.g. water or air) flowing through the exchanger also flow through these tiny holes. This causes the local formation of mini-streams to hit the heat exchange surface and increase the intensity of this exchange. Turbulence in the flow and rapid mixing of streams with each other play a significant role here.

What made you choose OpenFOAM software for numerical analysis? How demanding was working with the code?

At a particular stage of work on my doctorate, I noticed the possibilities and advantages of open-source software (e.g. OpenFOAM). Indeed, the beginnings of work in this program were tough. However, after half a year of scientific internship at TU Delft in the Netherlands, where I had the opportunity to work with a team with some OpenFOAM experience, I mastered the basics. Then I started to extend the possibilities of the program with my modifications. As a result, I implemented two previously unavailable advanced models of turbulence, which adequately represented the thermal and flow phenomena during the operation of the exchanger.

How did you rely on the resources provided by Cyfronet in this regard?

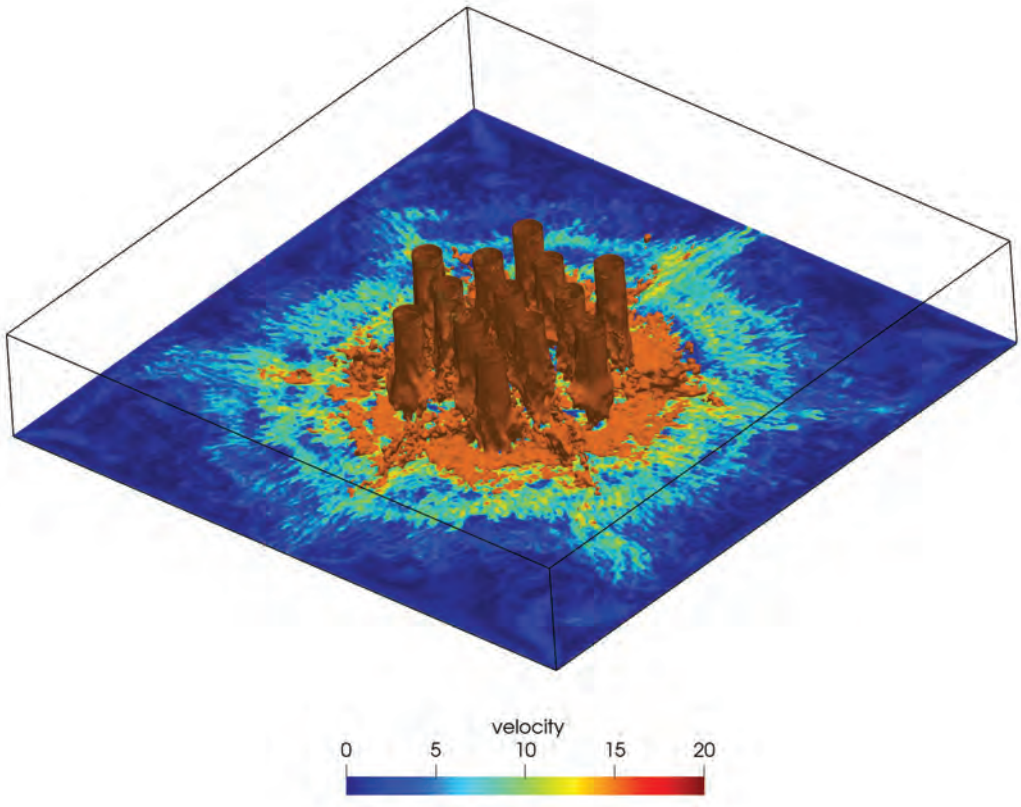
Virtually all my calculations and the programming environment were based on Cyfronet supercomputers (first Zeus, then Prometheus). The modifications I made were related to compiling my scripts, which required a properly prepared environment. In turn, the calculations themselves required a non-stationary approach at a particular stage (that is, considering the phenomenon's dynamics). Such simulations are extremely time-consuming – and when it comes to modelling turbulence, also resource-intensive. The most extended simulations lasted about a week, using, for

example, 64 computing cores simultaneously. I could not finish my doctorate without Cyfronet and only with my laptop.

What advice would you share with people just starting or planning to start research on their way to a PhD?

Indeed, you should not be discouraged by failures, especially at the beginning. It would be best if you always had your goal in mind. And it is imperative that you write down all your ideas and keep those notes in check! Several times during the preparation of the work, I found that at a given stage, I was implementing ideas that I had thought about much earlier, but I forgot about them.

Thank you.



Velocity isocontours of 13 jets impinging the flat surface. LES based analysis

Agnieszka Ogrodnik

The interview with the author of the PhD thesis:
"Measurement of photon-induced processes in heavy-ion collisions with the ATLAS detector"

Studying at the AGH University of Science and Technology was a springboard for you to cooperate with scientists at the international level...

That's true. Issues in the field of particle physics – a science that tries to answer basic questions about how our world was created and structured – interested me even in high school. I was studying Technical Physics at AGH. After my sixth semester, the best students had a chance to have international internships. I was offered a 7 week-long internship at the European Organization for Nuclear Research, CERN. I became a member of the ATLAS Collaboration. During my internship, I was working on a project that became the topic of my bachelor's thesis. I liked working with my supervisor, Prof. Iwona Grabowska-Bołd, we were getting along well, so I started another project with her on my master's studies and decided to start my PhD studies. I had the opportunity to work in an international group. The exciting thing is that I make a small contribution to a better understanding of the world surrounding us.

You mentioned that you deal with the physics of elementary particles. What exactly is your dissertation about?

In my PhD thesis, I studied two processes which have been theoretically described for many years. One has never been observed directly before, and the other has yet to be studied in detail in the measured kinematic range. Measurements in the field of particle physics are very commonly using huge accelerators to collect the data. Currently, the largest accelerator is the Large Hadron Collider located at CERN, enabling collisions of protons and lead-nuclei beams. Interestingly, the relativistic ions are surrounded by the electromagnetic field, which can be considered as a beam of photons. Due to that fact, apart from the strong interactions between protons and neutrons, we also observe photon-induced processes. The latter becomes dominant in ultraperipheral collisions when the distance between interacting nuclei is greater than twice the nuclear radius.

As I said before, I analysed two processes in the doctoral thesis: photon-photon scattering and the main source for this process, the exclusive production of the electron-positron pairs. The photon-photon scattering is interesting as a fundamental but very rare effect of quantum electrodynamics. It can be considered as a background for the processes beyond the Standard Model. Exclusive production of the electron-positron pairs is also one of the fundamental processes in quantum electrodynamics, and its measurement is essential to validate the modelling of the photon-induced processes. Additionally, it is a reference measurement for other analyses using the ultraperipheral collision data. Comparison with a reference enables improvement of the measurement precision by limiting the impact of systematic uncertainties.

The description of the exclusive production of electron-positron pair measurement was the main focus of my thesis. The analysed data was from lead-lead collisions recorded with the ATLAS detector at the LHC. The energy of the final state particles was low (of the order of a few GeV) and close to the detector electronics noise, which posed an experimental challenge. Part of the task was to optimise the selection used in real time to record the collisions with either of the signal processes. The high efficiency of this selection enabled the recording of a large number of events, which resulted in a decrease in the statistical uncertainty. The low efficiency of reconstruction and identification of particles with the energy of a few GeV is a bottleneck for this kind of measurement.

The obtained results were consistent with several predictions based on the Standard Model. Minor observed discrepancies suggest that the modelling of the photon-induced processes could be improved.

Big data requires "big" computers...

Definitely! I was using two Cyfronet supercomputers: Zeus and Prometheus. In addition, the ATLAS experiment software was required. I also needed to use some new software for simulation, and Cyfronet's Helpdesk expert quickly installed that. Any measurement in particle physics requires the multiple analysis of a large amount of data. Although it is possible to use the global GRID network, direct access to computing machines significantly speeds up work. Experience with the use of the queuing system is also helpful in working on supercomputers in institutes around the world.

How do you remember your doctoral studies?

I have perfect memories of the work with my team during my PhD studies. I have learnt a lot, and I do not only mean scientifically, but also in time management and task planning. I have two sons, aged 4 and 6, so frequent trips have been a big challenge for me (and my family!) Currently, I work at Charles University in Prague, where I continue my research in the heavy-ion group. I participate in the preparation of the trigger system for the ion--ion collision data taking that will take place this autumn.

If someone is considering starting the PhD studies or even master studies, I think it is important to focus on the topic you want to study and get to know the people you will be working with. Several years of work on one topic is quite long. Moments of exhaustion or discouragement may come. I don't think there is one perfect way to overcome it, but focusing on what we have achieved so far is always good. Also, try to set a way to your goal in baby steps. Then, you will also enjoy the little successes and see the progress you make.

Thank you for your time.

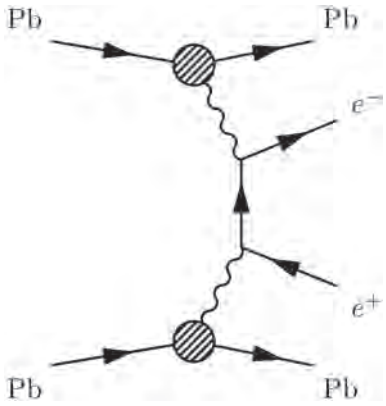
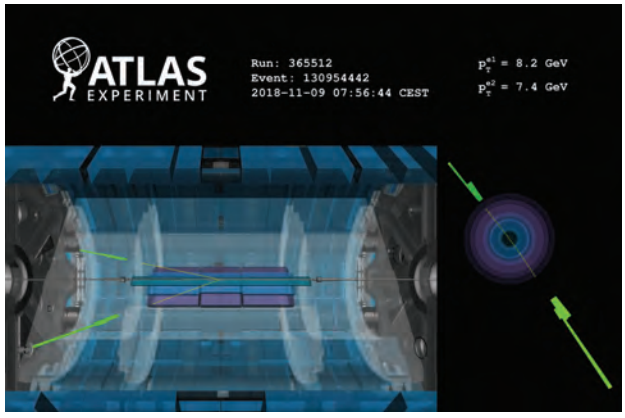


Diagram of the exclusive production of electron-positron pairs. Photons (wavy lines) originating from lead ions interact and create an electron-positron pair (straight lines). The ions stay intact



Visualisation of the exclusive production of electron-positron pair in the ATLAS detector's ultraperipheral collision of lead ions. The yellow lines mark the tracks left by electrons and positrons in the tracking detector. Marked in green are elements of the calorimeter, where these particles are stopped, and their energy is measured. [https://cds.cern.ch/record/2646921]

Tomasz Skóra

The interview with the author of the PhD thesis:
"Diffusion and reactions under crowding: Theory and simulations"

When did you decide to devote your time to science?

I went to the university already intending to become a scientist, although at the time, I was unaware of what that meant. Since childhood, I have been interested in science, especially chemistry. Probably the books from the cult series "Horrible Science" and in particular "Chemical Chaos" played a significant role in this. A slightly shabby copy still occupies an important place in my book collection, and I return to it from time to time.

While studying at the University of Silesia, I was following my plan and developing my scientific interests. I became involved in the activities of the "Aqua Regia" – Chemists' Scientific Circle. I moved to Jagiellonian University for my master's studies, where I quickly found a place in the Department of Theoretical Chemistry. I was introduced to the scientific craft there by Prof. Petelenz, who had the greatest influence on my idea of the work and ethos of a scientist. However, I decided to change my research subject and moved to the Institute of Physical Chemistry of the Polish Academy of Sciences, where I started in a slightly different research field on the frontier of biology, physics and chemistry. My advisor, Dr. Kondrat, taught me many aspects of research and networking. The doctorate was stressful at times, but it brought me a lot of satisfaction. I recommend it to everyone curious about the world but also humble and patient because the discovery process can be arduous.

Was it challenging to choose the direction of further development?

In the early stage of my studies, I did not attach myself to any specific issue or research area. In some sense, the subject found me — it was determined by the grant that funded my doctoral studies: diffusion in crowded environments. However, this is a vast field of study, so I had a lot of freedom in choosing the specific problems that I would explore in my doctoral thesis and the research methods that I would use for this purpose.

Please briefly present what your dissertation is about.

One of the major challenges of 21st-century science is understanding how biological cells work in terms of their building blocks (proteins, DNA, RNA and others). The most fundamental activities of biomolecules are their motion within the cell and undergoing chemical reactions. While we have gained extensive knowledge of these processes by isolating biological substances and studying them under controlled conditions, there has been growing evidence that these processes in cells occur differently than in test tubes. This is due, among others, to the fact that the cells are extremely crowded with macromolecules. My goal was to investigate what happens under the influence of this "crowding".

For instance, the slowdown of motion due to the crowding of the system with spheres is already a well-understood phenomenon. I decided to check, using computer simulations, what happens if elongated particles, instead of spheres, occupy the same volume. It turns out that the mobility slowdown is then substantially greater. Our collaborators from the RWTH Aachen University observed a similar effect in fluorescence correlation spectroscopy studies. Furthermore, I showed that crowding could cause reactions to become cooperative, which in practice manifests itself in their on-off behaviour upon changes in certain conditions, such as temperature or ion concentration. Biology is full of such on-off reactions, and it turns out that this feature can also arise due to crowding.

What was necessary to run your simulations?

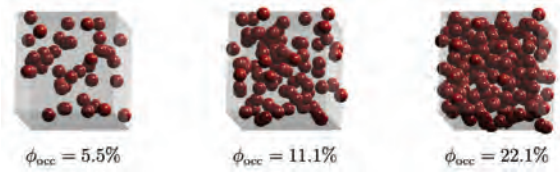
In the beginning, I used the Brownian dynamics package BD BOX. Later, I developed my own package for similar simulations, pyBrown, and performed subsequent simulations using it. In both cases, the Helpdesk Team from Cyfronet supported me when I encountered any problems. I could always count on help when I wasn't sure about something or when something "stubbornly refused to cooperate". I participated in a free training course that introduced me to the issues related to the use of high-performance computers. During my doctoral studies, I had the opportunity to use Prometheus, at that moment, the fastest computer in Poland, which was very exciting in itself.

While scientific progress is based more on ideas than on tools, it must be acknowledged that supercomputers are a tool without which research in a field like mine would be limited. Thanks to the PLGrid infrastructure access, I could ask much more challenging research questions and maintain the quality of calculations. In my research, on the one hand, a single simulation had to be extremely long to allow the moving particles to "feel" that they were in a crowded environment. On the other hand, dynamics of this type are stochastic, and each simulation yields different results. What we want to obtain, i.e., diffusion coefficient, results from averaging the random motions of thousands of molecules. In addition, each simulation had to be run many times – in many copies. I recommend Cyfronet research resources to anyone encountering similar technical challenges in their research work.

How do you remember your doctoral studies?

The work of a theorist is primarily solitary work. However, I had the pleasure of working with many different scientists. I owe it mainly to my advisor, who put great emphasis on research collaboration, for which I am very grateful to him.

The feeling of understanding brought the most joy. It's the feeling that the world has suddenly become less obscure somehow. Unfortunately, this feeling is rather sparingly and irregularly dosed. Hence patience and humility are recommended. Doctoral studies are not "just" more difficult than master's studies. It's not about grades, solving exercises, etc.; the main challenge is asking good questions. It's worth knowing this to save yourself the initial feeling of being lost.



Biological systems are characterised by the so-called macromolecular crowding, which has a large impact on the mobility of molecules and the equilibria of chemical reactions. In my work, I studied these effects using computer simulations, representing macromolecules with simple models, such as spheres and chains of spheres, of different concentrations

What are your plans for the future?

The next stage of my career is a postdoc at the University of Utah in Salt Lake City. I plan to make the most of this time to start an independent research career after returning to Poland. Like footballers, I only think about the next game. When the time comes to take the next step in my career, I hope to be ready for this.

Thank you for the interview, and we wish you further success.

Ewa Szymanek

The interview with the author of the PhD thesis:
"Numerical modeling and analysis of heat and mass flow in granular materials"

How did your adventure in Engineering begin?

I am a mathematician (Pedagogical University of Krakow), but I have always been interested in application aspects. Right after graduation, I joined the first research group (Czestochowa University of Technology) and I started dealing with the issue of granular materials. Then there was the moment of looking for "my place in science" and the decision to extend the research on granular materials (apart from the experiment and mathematical models) with numerical simulations.

Could you give us an idea of the importance of the issue that is the subject of the work?

The PhD thesis was devoted to the modeling and analysis of heat and mass flow processes through beds of granular materials and over surfaces of relatively complex geometry. These are flow issues whose correct modeling requires the use of advanced numerical methods. Due to the complicated geometry of granular beds, the process of generating computational meshes is difficult and extremely time-consuming. To overcome this difficulty, a computational algorithm was developed that combines the Immersed Boundary (IB) method with a computational algorithm dedicated to low-Mach number flow analyses, where spatial discretization is performed using high-order compact methods on partially shifted grids. The use of the IB method made it possible to perform simulations on Cartesian meshes, which eliminated the need to generate computational meshes adapted to solid surfaces.

Will a properly working model have an impact on engineering practice?

Definitely yes. Unlike commercial or typical finite-volume computational programs that mainly use second-order approximation schemes, the discretization method used in this work was based on sixth-order compact schemes and fifth-order WENO schemes. Consequently, even if a decrease in solution accuracy near the solid surface is unavoidable with the IB-VP approach, higher-order schemes allow for high accuracy far from solid objects. A weakness of higher-order methods is that they become unstable when the flow domain is locally discontinuous. However, it was shown that the formulated algorithm was stable in various cases, even in configurations with densely packed beds and large temperature differences. This behavior is attributed to the stabilising effect obtained by shifting the pressure node and interpolating the velocity components. Undoubtedly, the proposed method will facilitate the work, if only because of the differences that affect the time of obtaining the solution. On the one hand, we have computational time, and on the other, the entire time cost of preparing a given computational case should be taken into account. Using the method proposed in the PhD thesis, preprocessing takes much less time, which can sometimes compensate for the slightly longer time of the simulation itself.

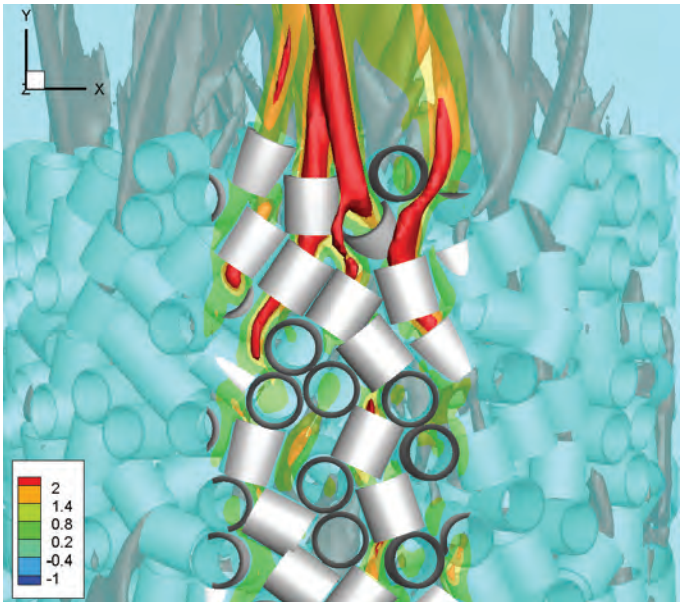
Would it have been possible to implement your research without relying on the Cyfronet's infrastructure?

Without the possibility of using Cyfronet's infrastructure, the implementation of my research would be impossible. As a member of the research team, I had a computational grant and was able to run simulations of the studied processes, which were very computationally demanding due to their complexity. Supercomputers make it possible to significantly shorten the calculation time, which in my case undoubtedly facilitated parallel experimental research. A big advantage is also the possibility of using training and professional and quick help in solving problems.

What was the single biggest challenge you encountered while completing your PhD?

Undoubtedly, it was a challenge to demonstrate that the applied immersed boundary method allows for the correct reflection of global and local changes in temperature and velocity caused by solid objects, taking into account the fine-scale flow structures in their immediate vicinity. At the same time, it was necessary to show that the computational algorithm created on the basis of a combination of the immersed edge method with the high-order compact discretization method, for which the solution's independence from the computational mesh density is obtained with a small number of nodes, is stable and effective.

Thank you.



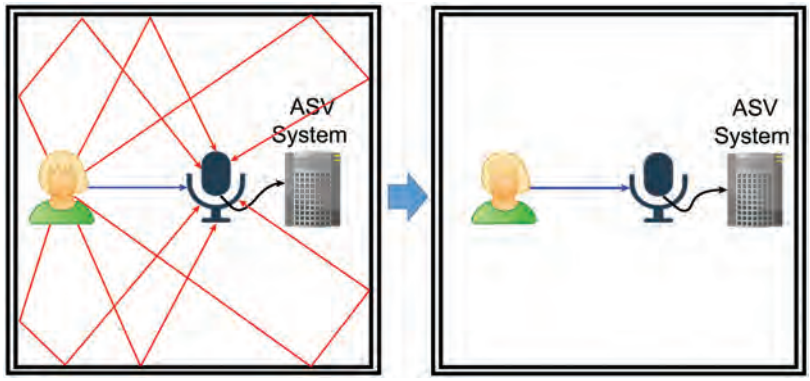
Axial velocity isosurfaces for bed elements temperature of 473 K and Reynolds number Re = 1000

Marcin Witkowski

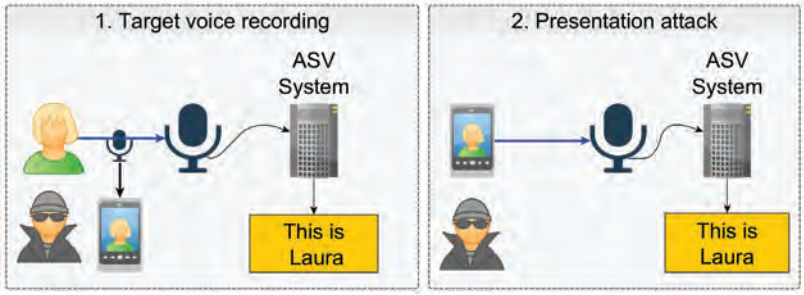
The interview with the author of the PhD thesis:
"Robust speaker verification with reverberation suppression and spoofing detection"

What made you interested in this particular aspect of security in telecommunication?

In my early years in the AGH Signal Processing Group at the Department of Electronics, I had the pleasure of working on a project devoted to the development of a speaker verification system. At this time, I noticed the basic scientific problems affecting the deployment of this technology, which outlined my research path. Over the past 15 years, we have seen two breakthroughs in speaker recognition methods based solely on the audio signal. In general, the solutions achieve satisfactory performance when benchmarked with typical audio datasets, but when the acoustic conditions during user recording are different from those during verification, the accuracy of a system drops significantly. Another aspect considered during deployment is the security of the system considered as its ability to automatic attack detection. In my work, I addressed both problems. First, I have proposed a method for the detection of presentation attacks, i.e. attacks in which a user's voice is recorded and then played back by an attacker to gain access to a service. I have also proposed a method that reduces reverberation from an already recorded audio signal, which might be used to minimise differences in acoustic conditions between the recording and the test. All in all, the goal of my dissertation was to develop solutions which enhance the application of voice biometrics technology.



Dereverberation



The presentation attack diagram

Do the recent successes in the field of publicly available AI platforms affect the possibilities of attack as well as defense against it?

In the last couple of years, there has been a significant increase in research works aiming at more accurate speech reproduction methods and algorithms which allow for the detection that an audio signal was synthesised or used in a spoofing attack on the speaker verification system. The effectiveness of attack detection can be tracked from publications from the Interspeech conference and, in particular, the proceedings related to ASVspoof or Voice Privacy Challenges. Note that both speech synthesis and voice conversion technologies based on deep neural networks already perform very well and allow for generating the speech of a person for whom a short audio utterance is available. I know a few AGH graduates who focus on this matter and achieve amazing results in their work. At this point, publicly available language models, often referred to as an artificial intelligence, rely on repositories with the source code for relatively old models (from 2 years ago) and generate the text in a particular person's style. However, I expect that the demand for voice communication with language models will indirectly accelerate the development of both voice conversion methods and protection against them.

How have Cyfronet resources contributed to the computing and analysis part of your research?

In my work I have used Cyfronet resources to parallelise calculations with and without use of graphics cards. In particular, in experiments devoted to developing a reverb suppression method, I used the Prometheus supercomputer to remove reverberation and evaluate the dereveberation performance objectively. Parallelisation involved processing multiple audio files simultaneously. The GPGPUs were used to train models for speaker features' extraction (speaker embedding) and further for speaker verification evaluation. Lack of access to the PLGrid computing infrastructure would have resulted in increased computation time of several to hundreds of times depending on the experiment, which would have definitely thwarted me from publishing my research results at significant conferences and completing the research for my thesis.

What could you advise people who are just beginning their doctoral studies? What should they pay the most attention to?

Scientific research is inherently difficult, as it requires navigating unfamiliar topics, where answers must be sought independently but not alone. A supervisor and access to a research infrastructure have a direct influence on success or failure. A good relationship with the supervisor, who is able to assist with advice or constructive conversation in the research area, is in my opinion crucial. In addition, when selecting a research scope, one should consider, if he/she will have access to a necessary computing infrastructure. For example, if you want to focus on machine learning, it is essential to have access to data as well as machines, such as those provided by Cyfronet that allow for large-scale experiments.

Thank you for the interview.

Photo: Anna Grudzień

Przemysław Zaręba

The interview with the author of the PhD thesis: "Synthesis, properties and conformational analysis of the new N-hexylarylpiperazines as ligands for aminergic receptors"

How did your scientific interests develop?

Already in high school, I wanted to search for new drugs. Studies at the Cracow University of Technology allowed me to fulfil myself in this area and work on the technological aspect of producing active substances. At this stage, I started research related to ligands of serotonin and dopamine receptors from the group of long-chain aryl piperazines, continuing them as a contractor in the NCBiR Lider project "Synthesis and evaluation of the activity of new ligands acting on the central nervous system (CNS)" (head: Jolanta Jaśkowska, PhD Eng.). During my doctoral studies, I became interested in computer methods in the design of biologically active molecules. I proposed a new low-base serotonin receptor ligands core using molecular modelling, which I then tested for anticancer activity. The molecules showed activity towards a glioblastoma cell line. Currently, I am conducting further work in this direction as part of the NCN Preludium project "New 5-HT6 receptor ligands from the group of sulfonamide derivatives of cyclic aryl guanidines in the treatment of glioblastoma multiforme", in which I am the principal investigator.

What made you choose N-hexylarylpiperazines as your research subject?

Long-chain aryl piperazines are an exciting group of molecules with a wide range of applications in the pharmacotherapy of central nervous system diseases. Some of them are successfully used to treat depression, anxiety or psychotic states. The rich spectrum of biological activity and the possibility of structural modifications prompted me to work with this core. A characteristic feature of the compounds I designed is an elongated carbon linker compared to typical long-chain aryl piperazines. This type of modification usually results in a change in the receptor affinity profile, which may lead to a different pharmacological effect. Some of the molecules I developed were subjected to ADMET tests to determine their pharmacokinetic properties and potential hepatotoxicity. Some ligands have been tested in in vivo behavioural trials, which allowed to confirm their antidepressant activity.

A group of new aryl piperazine derivatives being studied and developed by you became the subject of a patent. Can you explain what practical (production) implementations may result from that?

The patented results concern both the new, ecological method of *N*-hexylarylpiperazine synthesis in microwave radiation and the results of receptor affinity studies. The production method allows for a significant reduction in the reaction time: from several dozen hours to less than a minute. What's more, the developed solution is characterised by the high efficiency of the products obtained, the possibility of complete elimination or reduction of the amount of toxic solvents used or replacing them with more friendly ones (e.g. water), as well as the elimination of the need to use strong alkaline agents. These features are associated with a significant decrease in environmental nuisance and reduced production costs. The universal character of the method allows it to be used in obtaining a wide range of products. The implementation may reduce production costs, which in turn may reduce the drug's price. The second product covered by this patent application is a group of new *N*-hexylarylpiperazines with a characterised receptor affinity profile. In subsequent in vivo studies, the antidepressant efficacy of molecules belonging to this

group was confirmed. The patented library of compounds requires further research; however, based on the work carried out so far, the potential use of selected ligands in treating depression or psychotic states can be predicted.

What were the biggest challenges you encountered during your research, and how did you overcome them?

My research is characterised by high multidisciplinarity, covering aspects of broadly understood chemistry, pharmacy, biology and computer methods. The biggest challenge for me was to navigate within all these areas at the same time. Considering that I am a graduate of studies in chemical technology, it cost me much effort to get used to computer methods and their combination with pharmacology and molecular biology. Perseverance allowed me to overcome the problems I encountered, despite many setbacks. The help I received from other scientists from my university and other research units was also significant.

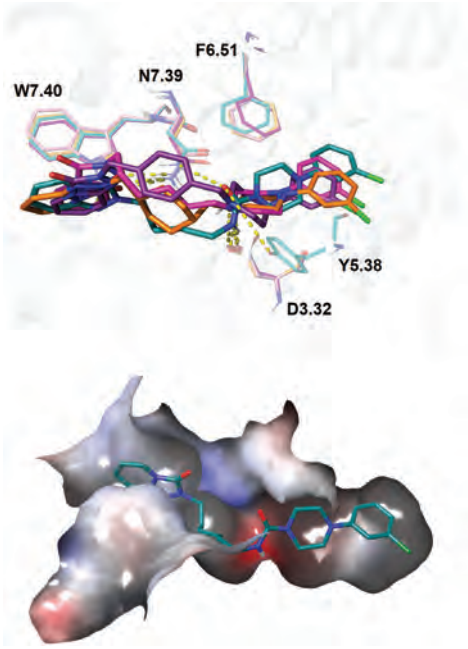
How did you use the IT resources provided by Cyfronet? What has been most helpful?

Using the computing infrastructure of Cyfronet, I applied molecular modelling in the design of biologically active compounds. A wide range of predictive methods, including homologous modelling, protein-ligand docking, QM/MM hybrid methods, quantum chemical calculations (PIEDA-FMO) or molecular dynamics, allowed me to precisely determine the structure of ligands for synthesis and to explain some structure-activity relationships. This type of approach led to a significant narrowing of the search area for active molecules, which in turn made it possible to shorten the duration of the research. Moreover, rejecting potentially inactive compounds is associated with economic and ecological benefits, which is particularly important considering the trend of green chemistry and sustainable scientific development.

What, in your opinion, should be remembered by young scientists who are taking their first steps on the research path?

In my opinion, the most essential thing in scientific work is consistency and perseverance. Every scientist encounters numerous problems on his/her way, both research and personal. However, persistence in pursuing the goal usually allows you to overcome difficulties. It is also important to be able to use the knowledge and experience of other scientists. Another important aspect is the reliability of the research and the presented results. We live in a world filled with false information. As scientists, we should be mainly involved in the fight against them and ensure that our data is accurate and supported by reliable research.

Thank you for your time.



Binding modes of N-hexylarylpiperazines into 5-HT1A receptor and matching to receptor binding pocket

Awards of the National Science Centre for users of ACC Cyfronet AGH computing resources

The award of the National Science Centre (NSC) is a distinction for young scientists for their significant scientific achievements, performed as part of basic research carried out in a Polish scientific unit, documented by publications affiliated to such a unit. The award was established by the Council of the National Science Centre in February 2013. The award is granted in three research areas: 1) humanities, social and art sciences, 2) life sciences, 3) technical sciences. Winners of these awards often emphasize the role of Cyfronet computing resources in achieving research results.

Award of the National Science Centre for Piotr Wcisło, PhD DSc from Nicolaus Copernicus University in Toruń*

On October 12, 2022, the laureate of the Prize in the category of technical sciences was the physicist Piotr Wcisło, a professor at the University of Nicolaus Copernicus in Toruń. The prize was awarded for the development of a new dark matter search method using optical atomic clocks and the use of ultra-precise laser spectroscopy to test quantum theory and search for new physics beyond the standard model.

– The usual didactic process does not cover the basics of quantum physics, so what I do is abstract from the perspective of most recipients – says the laureate.

The researcher developed a new, experimental method of searching for dark matter using ultra-precise laser spectroscopy based on the technology of optical atomic clocks and encouraged other centres in the world, where optical atomic clocks operate, as in Toruń, to create a network of detectors of this type. He has already carried out the first observations in this global observatory with partners from the United States, France and Japan.

– I use modern laser systems, combined with other technologies, like vacuum and cryogenic, to study the quantum theory. Using the capabilities of modern lasers, we observe the structure of molecules or atoms in ultra-precise detail – explains Prof. Piotr Wcisło. In his laboratory, a completely new laser system is being developed, in which he transfers current spectroscopic technologies to the regime of deep cryogenic temperatures, which will allow for a dramatic improvement in the accuracy of measuring the structure of the hydrogen molecule.

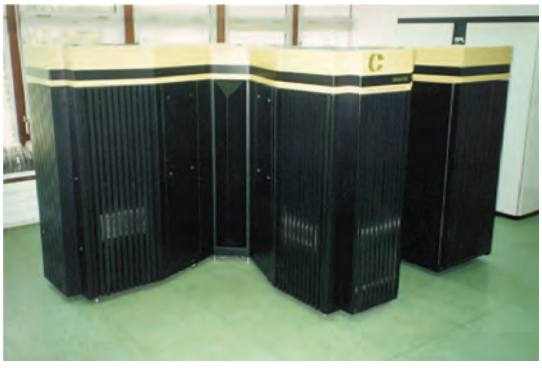
The scientist completed a postdoctoral fellowship at the American Joint Institute for Laboratory Astrophysics. He was also a Fulbright Scholar at the Harvard-Smithsonian Center for Astrophysics. He is the recipient of numerous awards and distinctions. He received, among others, Prime Minister's Award for his doctoral dissertation, the Award of the Minister of Science and Higher Education for scientific achievements, the scholarship of the Ministry of Science and Higher Education for outstanding young scientists, and the award for outstanding scientists of the Polish Academy of Sciences. He managed or is managing five NCN grants, projects of the Ministry of Science and Higher Education and NAWA, he was a scholarship holder of the FNP.

Prof. Piotr Wcisło supported his research with calculations on the Ares and Prometheus supercomputers at Cyfronet, within the *plgh2quantumtheory*: "Testing the quantum theory with molecular hydrogen" and *plgquantumscattering*: "Quantum scattering of diatomic molecules" computational grants.

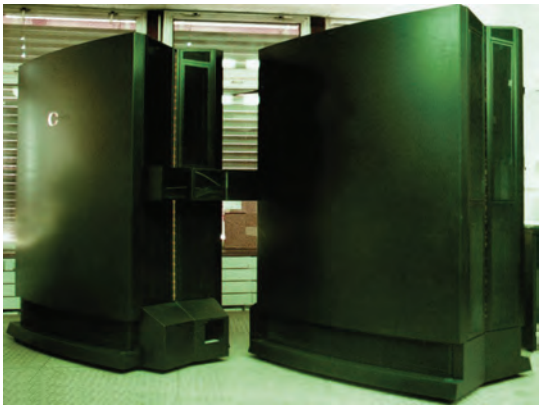
Congratulations to the winner and we wish him further success!

**Press material: https://www.ncn.gov.pl/*

CDC CYBER 72



Convex C3840



Exemplar SPP1600/XA

1973 CYFRONET is established

1975 A CDC CYBER 72 computer is deployed at the Centre

1990 The first Kraków node of the EARN / BITNET network is deployed at CYFRONET (on an IBM 4381 computer)

1991 CYFRONET installs a Convex 120 machine – the first vector computer in Central and Eastern Europe.

The first Polish national Internet link is established between Kraków and Warsaw.

Construction begins in the Kraków MAN

1994 A 2 Mbps link is deployed between Kraków and Warsaw

1996 An Exemplar SPP1600/XA computer deployed at CYFRONET took a position on the TOP500 list.

The first automatic tape library (ATL 2640) is installed at the Centre

1997 The ATM communications subnet is deployed within the Kraków MAN.

CYFRONET joins the POL-34 national backbone

1998 An SGI Origin2000 computer is deployed at the Centre

2000 Increasing the Centre network connection bandwidth to 155 Mbps

2002 A RackSaver PC computer is deployed at CYFRONET as part of the CrossGrid project

2003 An HP Integrity SuperDome computer is deployed at CYFRONET (the first such computer in Poland)

2005 An HP Storage Works XP12000 disk array is deployed at CYFRONET.

Increasing the Centre network connection bandwidth to 622 Mbps

2006 An HP Storage Works EVA 8000 disk array and an SGI ALTIX 3700 supercomputer (Baribal), with 0.8 TFlops of theoretical peak performance, is deployed at CYFRONET

2007 An agreement concerning the creation of the Polish Grid (PLGrid) Consortium was signed.

An SGI ALTIX 4700 supercomputer with the SGI RASC acceleration module is deployed at CYFRONET.

IBM BladeCenter HS21 servers are deployed at CYFRONET (6.2 TFlops).

An HP Storage Works EVA 8100 disk array is deployed at CYFRONET

2008 The configuration of SGI ALTIX 3700 supercomputer is extended to 1.5 TFlops.

MAN 10 Gbps started.

The Metropolitan Area Network is directly connected to Warsaw and Bielsko-Biała through the PIONIER network links, each of 2x10 Gbps capacity.

Zeus supercomputer (HP Cluster Platform 3000 BL) with 2 048 cores is deployed at CYFRONET

2009 Start of the PL-Grid project – Polish Infrastructure of Supporting Computational Science in the European Research Space

2010 The configuration of Zeus supercomputer is extended to 9,544 Intel Xeon cores, Zeus has been placed on 161st position on the TOP500 list

2011 Deployment of Hitachi Data Systems High Performance NAS Platform for computing infrastructure.

Total amount of installed disk space exceeds 2 PB.

The configuration of Zeus supercomputer is extended to 12,032 Intel Xeon cores.

Zeus has been placed on 80th position on the TOP500 list

2012 Start of the PLGrid Plus project – domain-oriented services and resources in the PL-Grid.

In April, ScaleMP, a leading provider of virtualisation solutions for high-end computing, announced that Zeus-vSMP system at CYFRONET is the largest virtual SMP system in Europe.

Zeus among 100 fastest supercomputers on the TOP500 list.

The Metropolitan Area Network is directly connected to Rzeszów through the PIONIER network link of 2x10 Gbps capacity

SGI Origin2000

SGI ALTIX 3700

HP Cluster Platform 3000 BL

Anniversary Medal

New Machine Hall

Prometheus supercomputer

2013 After upgrading of Zeus supercomputer configuration to 25,468 cores, its theoretical peak performance reached 374 TFlops.

Anniversary Medal has been minted

2014 The new Machine Hall is completed.

Start of two new projects – PLGrid NG and PLGrid Core.

The Metropolitan Area Network is directly connected to Katowice through the PIONIER network link of 2x10 Gbps capacity

2015 The Prometheus supercomputer (41,472 cores) is deployed at CYFRONET, and ranks high, 49th place on the TOP500 list (the July edition), and next (after upgrading to 53,568 cores) 38th place (the November edition).

For the first time in history two supercomputers from Cyfronet (Prometheus and Zeus) are ranked on the TOP500 list, in one edition.

The new backup Data Center is completed.

CYFRONET starts active participation in INDIGO-DataCloud, EGI-Engage, EPOS-IP and PRACE-4IP projects.

High Performance Computing centres in Poland (Gdańsk, Kraków, Poznań, Warsaw and Wrocław) are integrated with links of 2x100 Gbps capacity

2016 Prometheus ranks 48th (the June edition) and 59th place (the November edition) on the TOP500 list

2017 Prometheus ranks 71st (the June edition) and 77th place (the November edition) on the TOP500 list.

Further dynamic development of the Centre, including establishment of 6 new laboratories.

Sat4Envi, Gliomed, EPOS-PL and eXtreme DataCloud projects launched

2018 Prometheus (53,604 cores, 2.4 PFlops) ranks 103rd place (the June edition) and 131st (the November edition) on the TOP500 list.

EOSC-Hub and PRIMAGE projects have been launched

2019 Cyfronet represents Poland in the LUMI consortium, composed of eight countries that will jointly build one of the fastest European supercomputers.

Prometheus ranks 174th place (the June edition) and 241st place (the November edition) on the TOP500 list.

Cyfronet exhibition stand at the ISC'19 conference.

Cyfronet provides a new computational system for research using AI methods, with computational power over 4 PFlops for tensor operations and 256 TFlops for standard calculations.

PRACE-LAB, PRACE-6IP, SANO, EOSC-Synergy and EOSC Enhance projects launched

2020 Among the strategic infrastructures included in January 2020 on the Polish Research Infrastructure Map there are two projects proposed by ACC Cyfronet AGH as the initiator and coordinator of the PLGrid consortium: *National Supercomputing Infrastructure for EuroHPC* and *National Cloud Infrastructure PLGrid for EOSC*.

Prometheus (53,748 cores, 2.7 PFlops) ranks 288th place (the June edition) and 324th (the November edition) on the TOP500 list.

The Prometheus supercomputer supports scientists in the fight against coronavirus.

EPOS PL +, PRACE-LAB2, EPOS SP, PROTEUS-RS and EUROCC projects launched

2021 Ares supercomputer with theoretical peak performance over 4.0 PFlops is deployed in Cyfronet.

Prometheus ranks 373th and 440th place, and Ares ranks 216th and 267th place on the TOP500 list.

A new version of the ACC Cyfronet AGH website has been launched.

EuroHPC PL, PIONIER-LAB, KMD3, AGH – PANDA3, EGI ACE, EOSC Future and FINDR projects launched

2022 Athena supercomputer with theoretical peak performance 7.7 PFlops is deployed in Cyfronet.

Opening of the Data Center Podole.

For the first time in the history of Polish computer science, three supercomputers from one Polish computing centre (Athena, Ares and Prometheus) are ranked on the TOP500 list.

Prometheus (53,748 cores, 2.7 PFlops) ranks 475th place, Ares (37,824 cores, 4 PFlops) ranks 290th and 323rd place and Athena (6,144 cores, 7.7 PFlops) ranks 105th and 113th place on the TOP500 list.

Inauguration of the LUMI supercomputer.

EUMaster4HPC, FAIRCORE4EOSC, EuroScienceGateway, InterTwin, DT-GEO, Geo-INQUIRE and EDITH projects launched

2023 Celebration of the 50th anniversary of Cyfronet AGH with the participation of President Andrzej Duda, who presented state decorations to distinguished employees of Cyfronet.

Athena ranks 123rd place and Ares ranks 362nd place on the TOP500 list.

Cyfronet AGH honored with the Polonia Minor Award.

DOME, EUROCC 2, EUreka3D and PIONIER-Q projects launched

AGH UST Campus

LEGEND

1. Rector's Office
2. Faculty of Civil Engineering and Resource Management
3. Faculty of Metals Engineering and Industrial Computer Science
4. Faculty of Electrical Engineering, Automatics, Computer Science and Biomedical Engineering
5. Faculty of Computer Science, Electronics and Telecommunications
6. Faculty of Mechanical Engineering and Robotics
7. Faculty of Geology, Geophysics and Environmental Protection
8. Faculty of Mining Surveying and Environmental Engineering
9. Faculty of Materials Science and Ceramics
10. Faculty of Foundry Engineering
11. Faculty of Non-Ferrous Metals
12. Faculty of Drilling, Oil and Gas
13. Faculty of Management
14. Faculty of Energy and Fuels
15. Faculty of Physics and Applied Computer Science
16. Faculty of Applied Mathematics
17. Faculty of Humanities
18. AGH UST Academic Centre for Materials and Nanotechnology
19. AGH UST Centre of Energetics
20. Main Library
21. Department of Foreign Languages
22. Department of Sport and Physical Education
23. AGH UST Swimming Pool
24. Centre of e-Leaming
25. AGH UST Academic Computer Centre CYFRONET
26. Centre for IT Solutions
27. Education Centre
28. Centre for Student Affairs
29. AGH UST Doctoral School
30. Department for International Students
31. Admissions Centre
32. UNESCO Chair for Science, Technology and Engineering Education at AGH UST
33. AGH UST Student Campus
34. University Board of Student Government
35. Career Centre
36. Centre for Transfer of Technologies
37. Administration and Business Cooperation Department
38. Krakow Centre of Innovative Technologies INNOAGH
39. Centre for Project Management
40. Administrative Centre for Science
41. Space Technology Centre AGH UST
42. Department of International Relations
43. Disability Support Office
44. Geological Museum of the Faculty of Geology, Geophysics and Environmental Protection
45. AGH UST Press
46. Academic Cultural Centre Club STUDIO
47. Student Club Gwarek
48. Student Club Zaścianek
49. Student Club Filutek
50. Recording Studio Kotłownia

ISBN 978-83-61433-45-3

Academic Computer Centre
CYFRONET AGH
ul. Nawojki 11

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

cyfronet@cyfronet.pl
www.cyfronet.pl