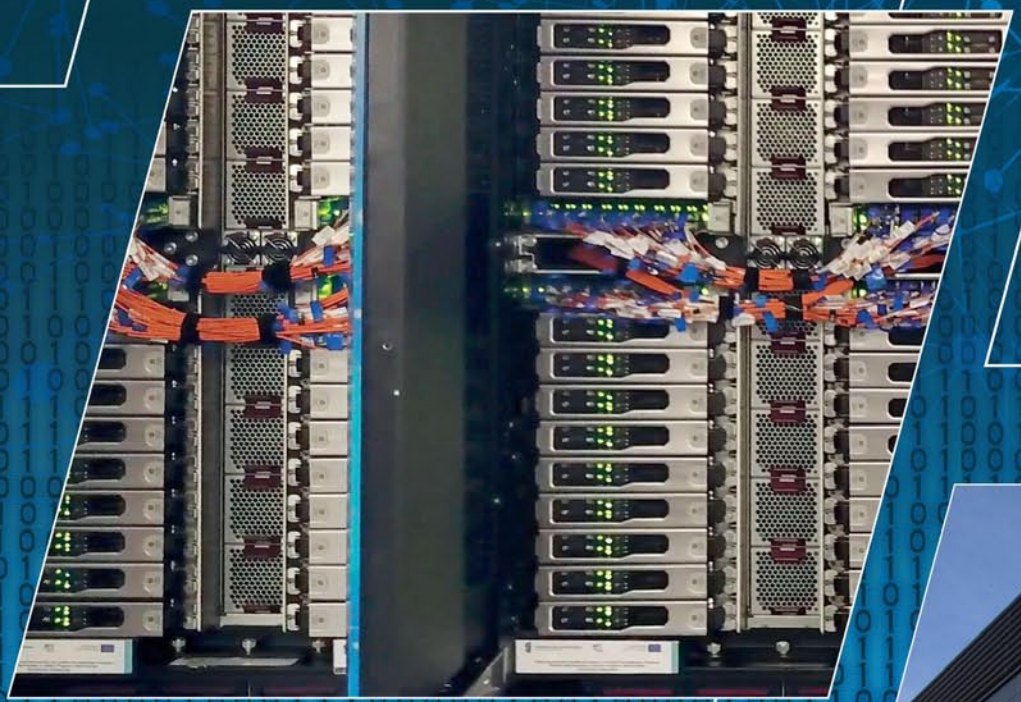
Academic Computer Centre

CYFRONET AGH



2022

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!

Another year burdened by the difficulties of the COVID-19 pandemic is behind us. Being more experienced, we look forward to the coming year with hope. In 2022, in addition to carrying out many daily tasks, Cyfronet will continue many works related to, inter alia, participation in LUMI and EuroHPC PL undertakings.

LUMI is a unique initiative involving ten European countries that are investing in a single supercomputing system with an enormous peak performance of 552 PetaFlops as part of the EuroHPC Joint Undertaking (EuroHPC JU). This is a big challenge for Cyfronet, as we have been tasked with organizing access for Polish scientists to one of the fastest and most ecological supercomputers in the world. LUMI aims to accelerate research and thus increase the competitiveness of European solutions. The supercomputer will support European scientists, industry and the public sector in meeting complex challenges, a.o. in research on medicine, biology, bioinfornatics, chemistry, quantum chemistry and molecular physics, high energy physics, new materials, nanotechnology, ecology, energy sector, meteorology and many, many others.

As part of the "National Supercomputing Infrastructure for EuroHPC - EuroHPC PL" project, Cyfronet, as the project coordinator and Consortium leader, is building, together with six Partners, a specialized general-purpose infrastructure for large-scale computing, which will enable undertaking research challenges in key areas from the point of view of Polish society, the scientific community and the economy. This infrastructure draws on the latest technologies in the field of High Performance Computing (HPC). Large-scale calculations carried out on this infrastructure, including simulations and analyses of large scientific data sets, are part of the EuroHPC JU, i.e. the European system of exascale supercomputers. Building and sharing computing resources under EuroHPC PL will significantly increase the scope and quality of data-based scientific research and the use of artificial intelligence methods in Poland.

Four supercomputers from Poland were present in the list of 500 fastest supercomputers in the world (TOP500 list, June 2021 edition), including two from Cyfronet: Ares and Prometheus, taking respectively 216th and 373rd place. Ares was launched in Cyfronet in 2021 and is built under the PraceLab project from a set of computing servers with over 3.5 PFlops of computing power and a disk subsystem with a capacity of over 11 PB. The supercomputer has 37,824 computing cores (Intel Xeon Platinum processors) and 147.7 TB of RAM. It is equipped with computationally efficient GP GPU cards and a liquid cooling system.

Scientific research carried out with the help of the Centre's IT resources concern many fields. We support researchers using our supercomputers in their scientific work and in making important discoveries. Each success of the Users of Cyfronet infrastructure is extremely important to us.

I would like to thank all our Friends and Users for cooperation and valuable advices regarding the further development of the Centre. I would like to invite you to personal contacts with Cyfronet employees. In the current situation, these will still be mainly remote contacts, but we hope that they will bring a lot of fruitful cooperation...

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 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

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

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

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

Technical Department Pychowice – 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 – Tomasz Szepieniec, MSc (manager)
Phone: +48 12 632 33 55, e-mail: Tomasz.Szepieniec@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

photo: Adam Frączek, KSAF AGH

photo: Krzysztof Mastalski, KSAF 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	Oil and Gas Institute National Research Institute
Marek Gorgoń, Prof.	AGH University of Science and Technology
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.	Pedagogical University in Cracow
Jerzy R. Jaworowski, PhD Eng.	Cracow University of Technology
Marek Kisiel-Dorohinicki, Prof.	AGH University of Science and Technology
Mariusz Kwinta-Pudełko, Msc Eng.	Agricultural University in Cracow
Wojciech Majka, MSc Eng.	Jagiellonian University Medical College
Tomasz Masier, MSc Eng.	Institute of Forensic Expertise
Roman Mazur SDB, PhD	Pontifical University of John Paul II in Cracow
Paweł Mikłaszewski, MSc	Academy of Music Academy of Fine Arts Academy of Theatre Arts
Wojciech Palacz, PhD	Jagiellonian University
Dariusz Put, Prof.	Cracow University of Economics
Elżbieta Tomczyk, MSc	Łukasiewicz Research Network Cracow Technological Institute
Jacek Barski, Deputy Director	Marshal Office of the Małopolska District
Paweł Schmidt, Director	Cracow City Hall

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:

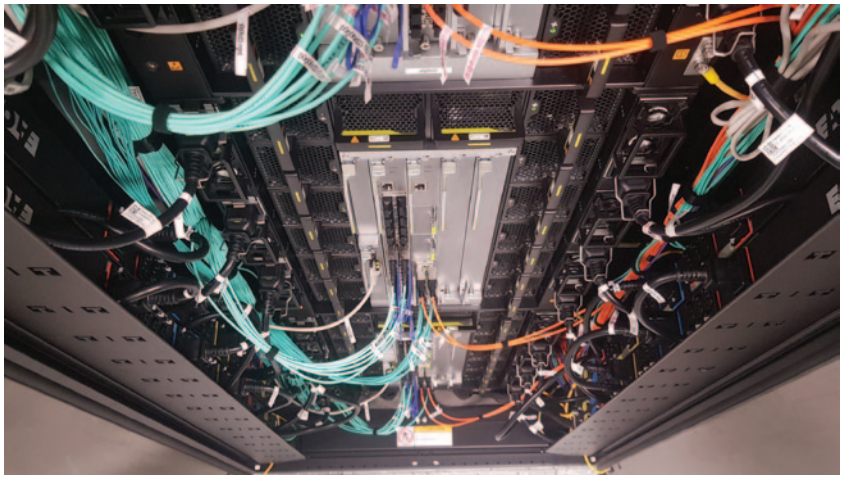
- 532 servers equipped with 192 GB of RAM,
- 256 servers with 384 GB of RAM,
- 9 servers, each with 8 NVIDIA Tesla V100 cards.

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 2021 edition)	216th position

The total theoretical performance of the CPU parts is over 3.5 PFlops, and of 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.

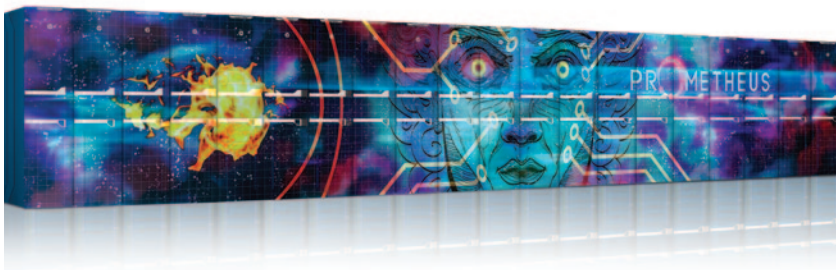
Ares complements Cyfronet's computing resources by providing a newer generation of processors and servers with more memory. This will enable shortening the computation time of scientific tasks and addressing problems that so far could not be run on a large scale due to insufficient memory. In addition,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


E 9000



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 – the most powerful Polish 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 2.65 PFlops (PetaFlops)!

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 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.

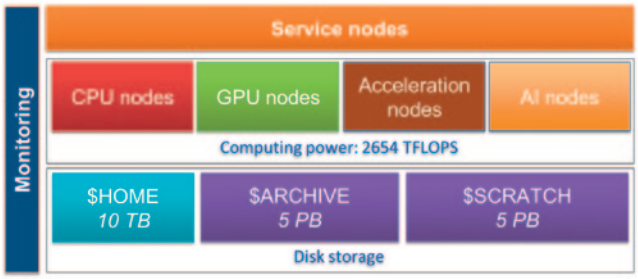
Prometheus in numbers	
Number of computing cores	53 748
RAM	283.5 TB
Number of GPGPUs	144
Computing power	2.65 PFlops
TOP500 – the list of the world's fastest computers (June 2021 edition)	373rd position

Prometheus once again has been listed on the TOP500 list of the world's fastest computers (June 2021 edition) and took the **373rd position**, as one of four supercomputers from Poland. The other supercomputer from Cyfronet, Ares, took 216th position.

Division into parts with diverse functionality, applied in the Zeus supercomputer, has been very well used by its users. Due to this fact, the Prometheus architecture is also 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 the fastest dedicated solution for artificial intelligence available for the needs of science in Poland.**

Prometheus architecture



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.

Year	No. of Jobs	CPU time in years
2015	1 099 822	5 811
2016	3 080 543	21 239
2017	5 032 438	36 600
2018	5 430 811	39 946
2019	2 738 534	41 829
2020	3 119 106	39 680

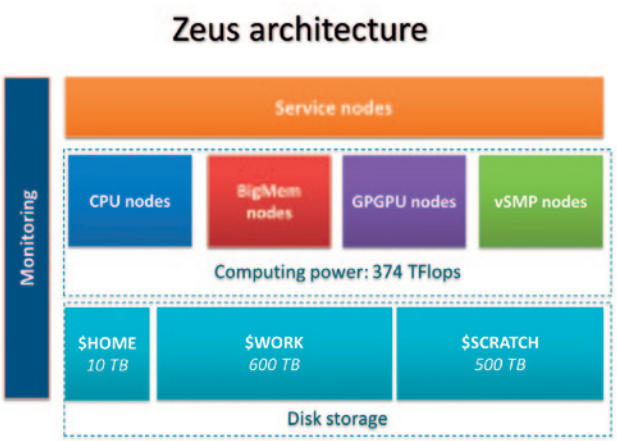
ZEUS – over 60 000 CPU-years

Cyfronet operates one of the fastest supercomputing systems in Poland, named *Zeus*. It currently provides 374 TFlops of theoretical performance, 25,468 CPU cores and over 200 GPGPUs. All this, equipped with 60 TB of RAM and 2.3 PB of disk storage supports the computations of scientific communities.

The Zeus supercomputer was launched in 2008. Since that time, it has been continually noted (12 times) on the TOP500 – the list of the world's fastest computers. Four of these locations were on **TOP100** subset, with **81 – the highest noted spot**. Zeus was 10 times the fastest in Poland.

The architecture

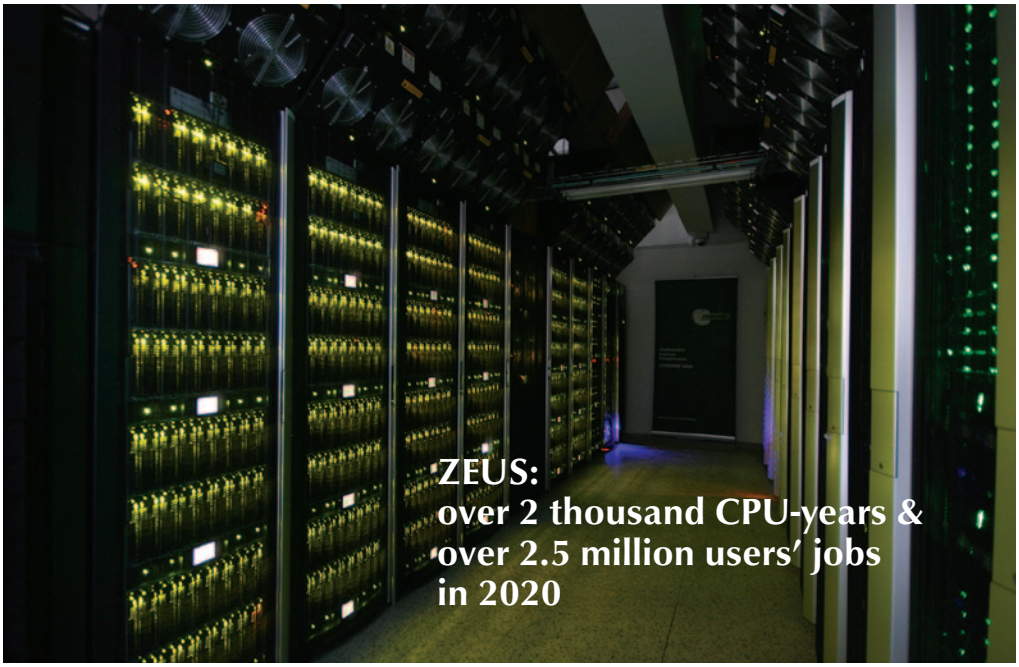
Zeus is a heterogeneous computing cluster. It constitutes of four classes of nodes, varying in terms of architecture of computing resources, specifically tailored to the requirements of the scientific communities. The Zeus architecture is a composite of four partitions:



- **classical cluster** of computing servers with highly efficient CPU nodes equipped with two Intel Xeon processors and 16-24 GB of memory per node,
- cluster of servers with large amount of memory – "**fat nodes**" with four AMD Opteron processors and 256 GB of memory per node,
- set of servers equipped with **GPGPU** accelerators (Intel Xeon processors as well as NVIDIA M2050 and NVIDIA M2090 cards) and **FPGA** accelerators (Pico Computing M-503 modules with Xilinx Virtex-6 LX240T),
- **"virtual" SMP** computer with large, shared memory, using vSMP software of the ScaleMP company – the nodes with Intel Xeon processors connected with a specialized virtual machine hypervisor, which allows for booting up the machines up to 768 cores and 6 TB of memory.

Diversification of the node types gives a possibility to fit users' applications to the hardware, which matches at best their characteristics and special requirements. For example, the classic CPU node group is dominated by serial and parallel (MPI) jobs, while the second one is great for large memory jobs. The GPU nodes allow some applications to benefit from GPGPU accelerators and the vSMP

nodes give a possibility to run huge memory jobs or scale applications, which do not use any inter-node communication library, like MPI, for parallelism. It is worth noting that Zeus-vSMP was the Europe's biggest installation of this type when launched!



The users

Since it was launched, the Zeus cluster has been serving the whole scientific community from Poland. In 2020 alone Zeus performed over **2.5 million tasks** serving scientists from many universities and research institutes. Majority of tasks were executed on more than one processor and this trend is constantly increasing. Typical user computations request even **several thousand of cores just for one job**!

Year	No. of Jobs	CPU time in years
2008	603 525	207
2009	2 227 804	876
2010	4 009 049	990
2011	7 557 817	5 052
2012	8 126 522	7 923
2013	7 932 978	11 016
2014	7 694 224	12 980
2015	6 405 941	10 141
2016	4 668 134	3 414
2017	4 034 454	2 632
2018	2 911 875	2 490
2019	2 255 105	2 198
2020	2 577 813	2 081

Supercomputers usage

Prometheus and Zeus are part of the European cloud and grid infrastructure under the European Grid Infrastructure (EGI). At the same time, Prometheus and Zeus 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 Prometheus and Zeus 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 Zeus and, recently, Prometheus is quite reach. It includes:

- dynamic analysis of ceramic composite materials,
- modeling and optimization of the prototype of an energy-saving hydrogen fuel cell vehicle,
- simulations of molecular dynamics of electrolytes for ion cells,
- use of quantum methods for image analysis,
- gene expression in cod from the Baltic Sea,
- study of the interaction of influenza virus fusion peptides with lipid membranes,
- analysis of the impact of underground fluid injection on drops of stress of induced seismic tremors,
- microphysical processes in space plasma,
- speech and handwriting processing and recognition.

A wide range of research topics is evidence of constantly increasing number of scientists, who are aware of advantages of supercomputers like Zeus or Prometheus. With the help of these powerful supercomputers 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.

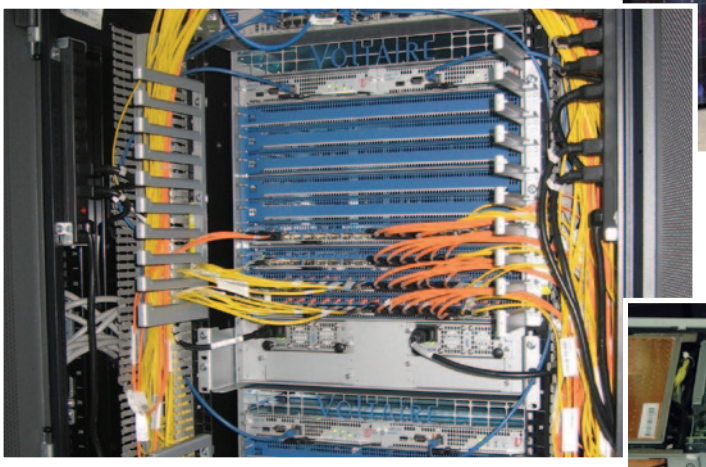
In addition to individual scientists and small research groups, even international consortia carry out calculations from many different scientific disciplines with the help of supercomputers of course with the participation of Polish scientists. Scientific computations do not include simulations only.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

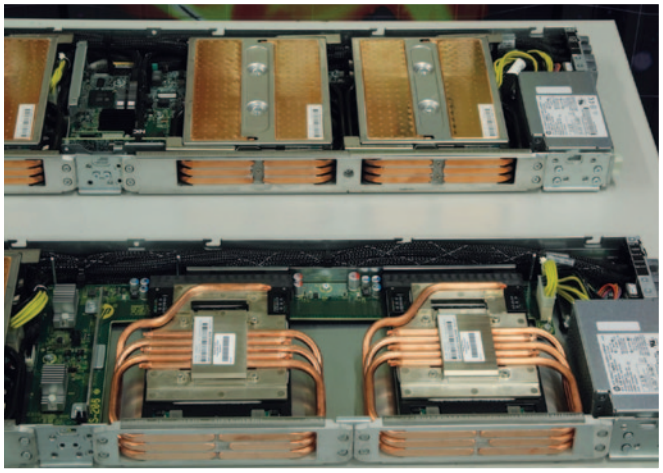
Obviously, even the highest positions in the TOP500 list, or the latest technologies used to build high-performance computers do not fully reflect the importance of this kind of computing resources for the Polish scientific community. The usefulness of supercomputers provided by ACC Cyfronet AGH as a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

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
Prometheus supercomputer		
2015	1 099 822	5 811
2016	3 080 543	21 239
2017	5 032 438	36 600
2018	5 430 811	39 946
2019	2 738 534	41 829
2020	3 119 106	39 680
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





Data storage

Only the proper teaming of computing infrastructure with the right selection of storage solutions can assure the best quality of services provided to users. The scale of problems in this area increases with the complexity and the efficiency of high performance computers. At present, disk storage systems attached to Cyfronet's supercomputers store over 500 000 000 data files (with the file sizes up to several terabytes). A wide variety of research conducted on the Centre's resources requires not only diverse configuration of key Cyfronet's computers, but also an efficient, dedicated storage system.

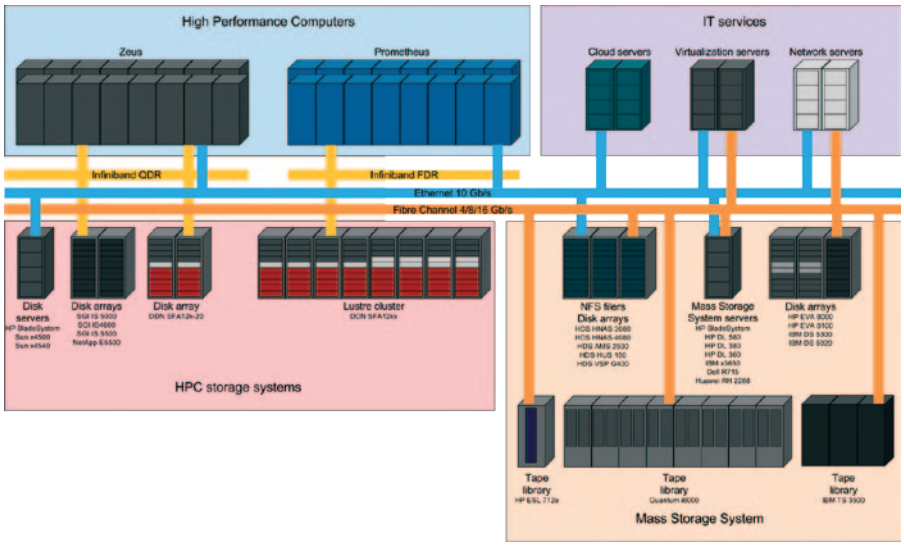
The most fundamental is the one used for keeping users' home directories. In this case all the crucial elements provide a very high level of availability and data security, which are supported by mechanisms such as snapshots and backups to external tape libraries. Zeus and Prometheus (the two main supercomputers of the Centre) offer such functionality through using specialized HNAS file servers (so called filers), produced by Hitachi Data Systems. These servers support hardware implementations of the NFS protocol and provide very high performance and high availability of the file systems. HNAS filers are coupled with Hitachi Data Systems AMS 2500 and HUS 150 disk arrays, used as repositories of physical disk space. These devices also provide extremely high levels of security and performance, fitted to the specific characteristics of the data stored in home directories.

Another type of storage space used in supercomputers is the scratch space, in which the crucial factor is speed. To address this requirement, Cyfronet uses the Lustre distributed file system, which is capable to scale both space and performance by aggregating storage capacity of many servers. Moreover, throughput and/or capacity can be easily increased by adding more servers dynamically, without interrupting user computations. Nowadays, all Cyfronet's supercomputers can use scratch spaces based on Lustre. In Zeus case, it is the file system with almost 600 TB capacity and 12 GB/s read/write bandwidth. Prometheus' scratch has enormous capacity of 5 PB and 120 GB/s read/write bandwidth. For even more demanding disk access requirements it is possible to use a super-fast RAM-disk provided by the vSMP partition of the Zeus supercomputer.

However, the major part of Cyfronet's storage resources is dedicated to the needs of users of domain-specific services developed in the PLGrid program. The PLGrid infrastructure provides a dedicated workspace for groups in domain grid environments – the functionality essential for enabling cooperation of scientists from geographically distributed locations. Zeus provides almost 200 TB of such disk space with the use of HNAS filers and the NFS protocol. Prometheus offers similar functionality with higher performance, using the Lustre file system. The maximum capacity of the */archive* resource in this supercomputer reaches 5 PB and the total rate of read/write operations attains 60 GB/s.

A special case of mass storage are the resources for large projects and international collaborations, in which Cyfronet takes part, such as WLCG (Worldwide LHC Computing Grid), which stores and analyzes the data coming out of the LHC detector in CERN, or CTA (Cherenkov Telescope Array). Such projects demand high volumes of disk space available by a set of specialized protocols, such as SRM, xroot or GridFTP. Cyfronet provides such space with the use of the DPM (Disk Pool Manager) instances and dedicated networks, such as LHCone. Total amount of disk space provided by these services exceeds 1 PB. The overall data storage space exceeds **59 PB**.

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 is clearly visible correlation between 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 resources is **34 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 6 thousand slots for LTO magnetic tape drives and 36 drives of the III, IV, V and VI generation. A single LTO-6 magnetic medium has a physical capacity of 2.5 TB and allows recording at the speed up to 160 MB/s, which theoretically allows the storage of almost 15 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 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.

For the special cases Cyfronet offers users a dedicated backup-archive service called the Universal Archiving. Within this service, the user is given a dedicated disk space, protected at many levels. User's data in this case is protected by a distributed disk array equipped with disk resources protected by RAID-6 level functionality and additionally secured by the HSM system of hierarchical data storage. To advance the data safety even more, users' backup data are additionally protected by geographical data replication to the associated units. Last but not least, users of Universal Archiving system can further increase their data safety by encrypting their data with use of certificates. At present, the total storage capacity of Cyfronet tape resources exceeds **25 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 consisting Interdisciplinary Centre for Mathematical and Computational Modelling in Warsaw, Poznan Supercomputing and Networking Center, Wroclaw Centre for Networking and Supercomputing, Academic Computer Centre in Gdansk and **Academic Computer Centre Cyfronet AGH as the initiator and coordinator 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 81st 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.

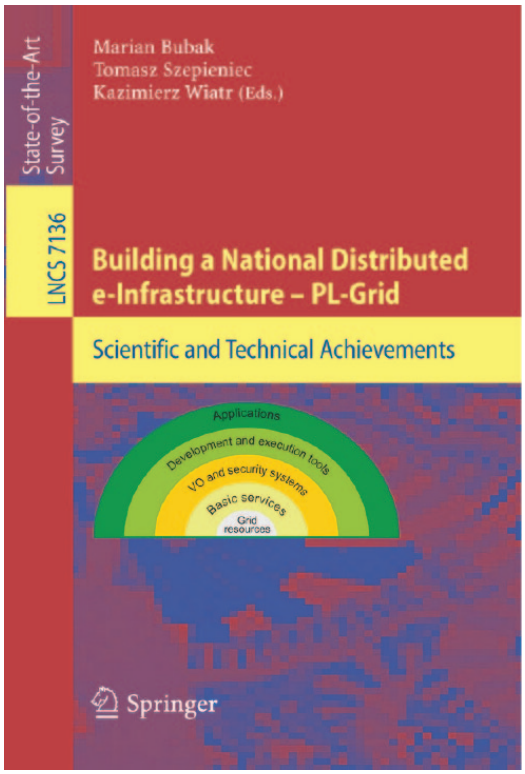
In the PLGrid NG project, the domain-specific grids were developed for several other groups of scientists, representing fourteen research fields (**in total, in the two projects, IT support tools were built for 27 scientific disciplines**).

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 recognized 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!**

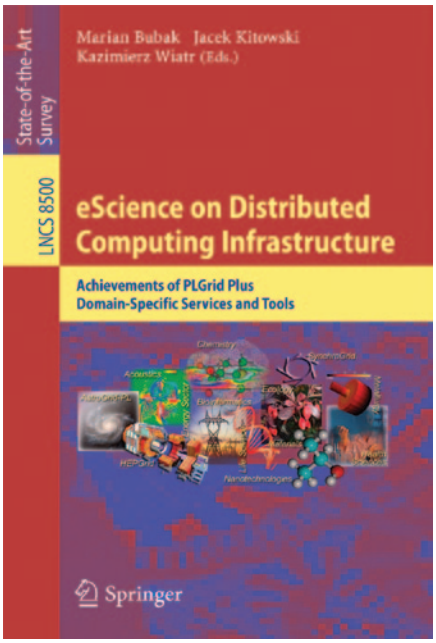
At present, more than 5 PFlops of computing power and more than 60 PB of disk storage are available within the infrastructure. In addition, many tools supporting organization of computational experiments, designing and running applications, computationally supporting research and results' visualization were implemented in the infrastructure. Furthermore, the Consortium introduced a new service – 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. Vast range of services contributes to increase of cooperation between Polish scientists and international groups of specialists from many different scientific domains – also humanities and social sciences. The essential fact is that anyone who is performing scientific research can be the user of the infrastructure. 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.



Domain-specific grids in the PLGrid infrastructure

The PLGrid infrastructure, established through the Cyfronet's initiative, offers a uniform access to resources of all five 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.

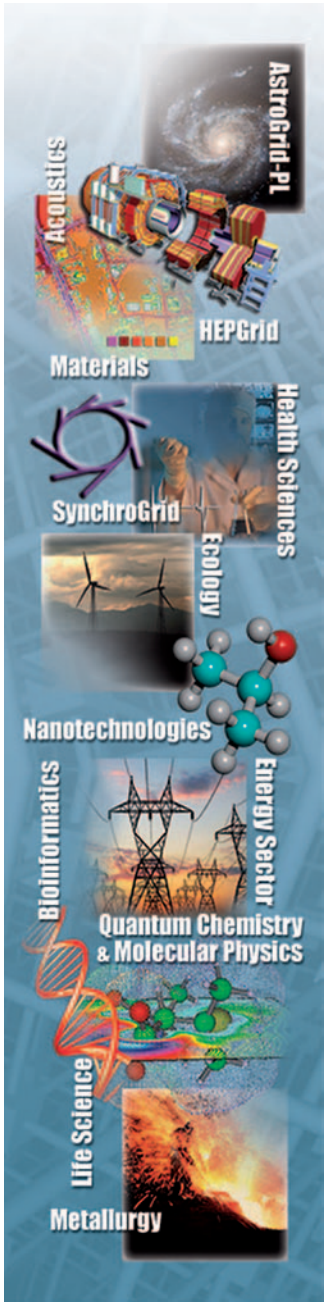
The second part (chapters 9 to 13) presents new environments and IT services that can be used by all of the previously mentioned groups of scientists. The third part (chapters 14 to 36) describes how specific environments, tools and services, prepared within the PLGrid Plus project, are used in advanced 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 specialized 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 specialized 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.

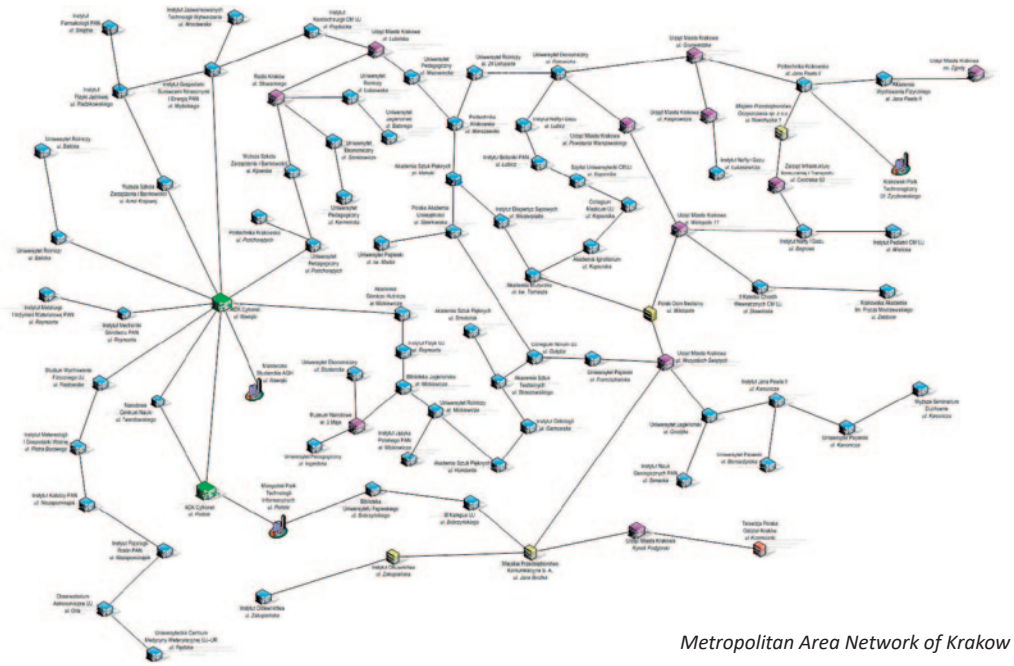
Preparation and implementation of a set of domain-specific services fit very well with the need of development of an advanced IT infrastructure designed for the implementation of modern scientific research. The well-tailored PLGrid e-infrastructure does not only fulfil researchers' needs for suitable 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.



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. In order 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 development and maintenance of the Metropolitan Area Network (MAN) to achieve its availability 24/7.



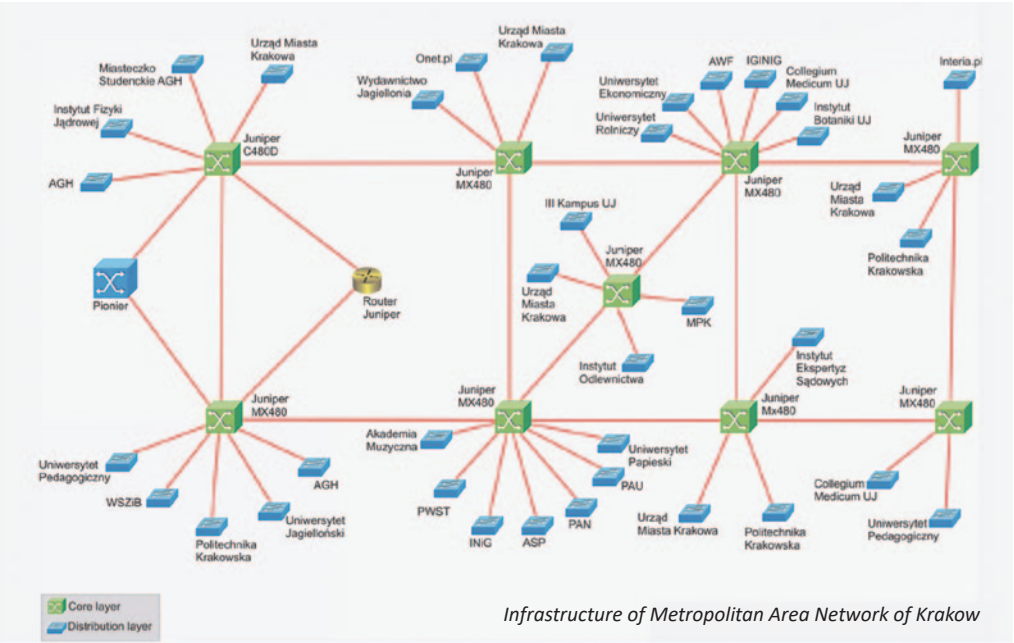
Metropolitan Area Network of Krakow

Main characteristics of MAN

It is not possible to attain high network availability without its constant development and adjustment to the needs of users. The length of dedicated fiber-optic links reached this year about 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 covers also Bronowice, Krowodrza, Czyżyny and Nowa Huta zones. 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 includes also other directions, up to the borders of Kraków.

The fiber-optic infrastructure is the basis of the MAN operation. ACC Cyfronet AGH takes 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 very important 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 to maintain the continuity of operation in situations 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 networks switches is connected with at least two and sometimes even three neighbours for automatic and transparent recovery in case of a failure of any network device or link. 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 enables also 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 CenturyLink Communications network.

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 specialized 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 on the world with a single authorized account, providing at all locations the same way as access to the network at the parent unit,
- **box**: a network drive (*http://box.cyfronet.pl*) allowing file exchange and synchronisation. The drive can be also accessed from mobile devices via dedicated application.

Network services in numbers in 2020	
Number of e-mails	~ 17 000 000
Number of e-mail server sesions	~ 45 000 000

Portals and mobile applications

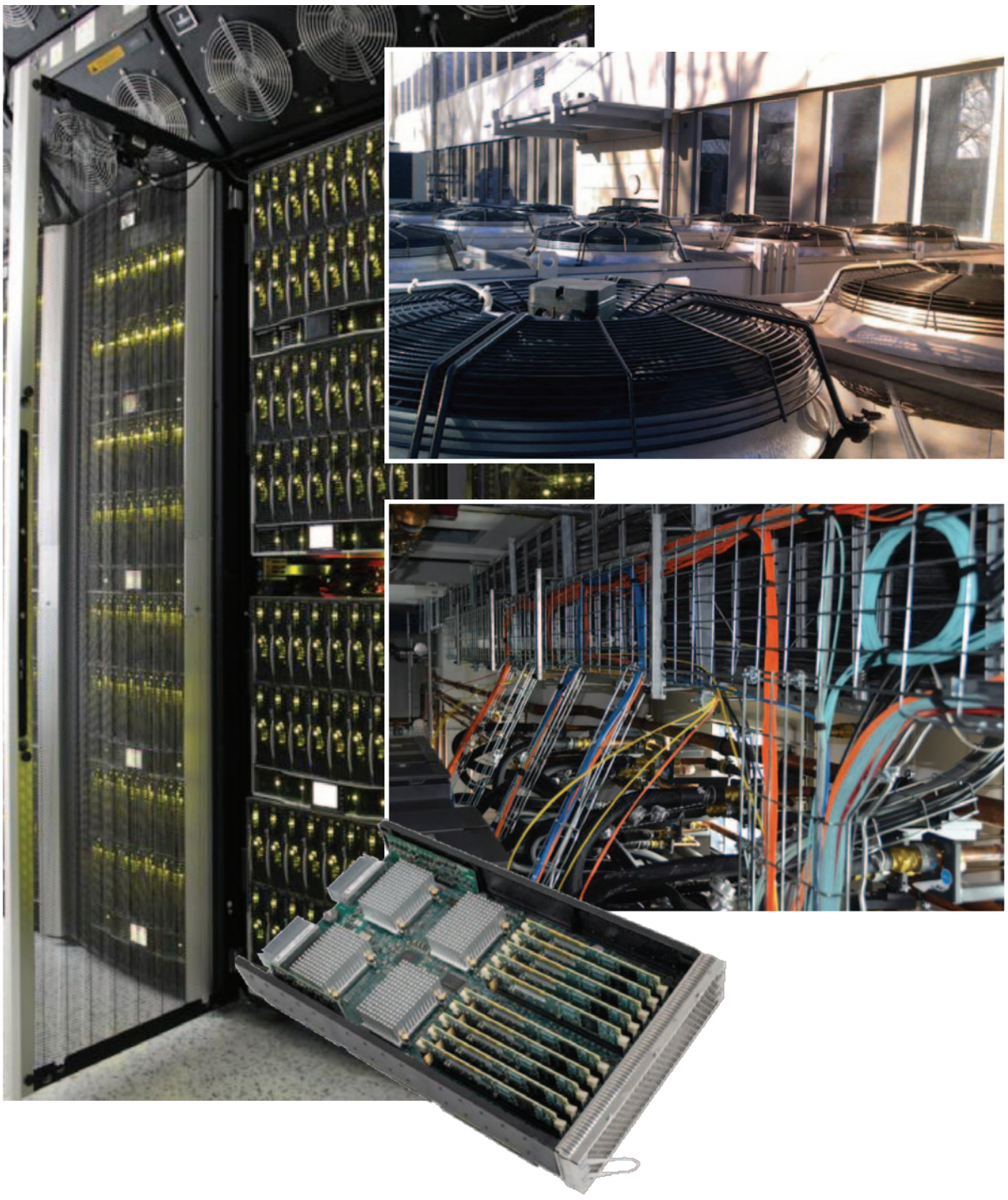
The Centre does not limit its activities to the scientific areas only – it also contributes to the development of the information society. The Web server at ACC Cyfronet AGH serves as an Internet hub for the entire Kraków scientific community. The Centre continues to develop and extend its Web portal, 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 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* has been 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 Polish, English and Spanish.



Computational resources

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

Computing resources

Prometheus, Zeus and Ares supercomputers provide:
7+ PFlops
117 000+ cores
450+ GPGPUs
500+ TB RAM

Storage

34 PB of disk and 25 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 Stash 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.

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 typical computing and storage platforms, we offer the PLGrid Cloud Computing Platform.

– *Up to now, the PLGrid infrastructure has been providing a set of well-defined environments with computing and storage resources. The cloud platform is not just an extension of them. We foresee it as a new quality level of conducting research* – says Kazimierz Wiatr, the Centre's Director. – *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* – adds Director.

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

- The Cloud increases elasticity of research, as scientists can tune the virtual machines to their specific needs. Up to now, to set-up a "virtual laboratory" solving some specific scientific problem, some help of PLGrid experts was needed. Now, each scientist can create and easily extend such virtual laboratory alone.
- 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. A start of a new VM takes just around 30 seconds.
- 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.

– *The Cloud Computing in PLGrid and Cyfronet is an innovative solution on a European scale. We have a strong belief it will bring a new quality level for research conducted by our users* – concludes Prof. Wiatr.

Currently **200+ various types of VMs** are utilised on Cyfronet resources.

ONEDATA

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

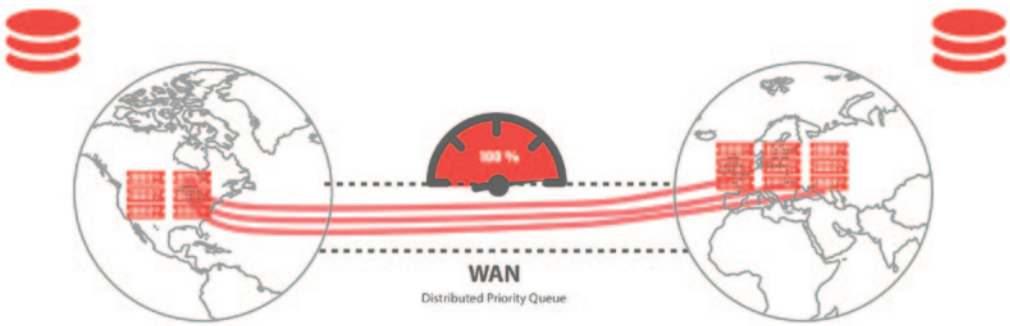
- Simple deployment based on Docker containers using a friendly command line client.
- Easy storage support for user requests based on secure tokens.
- Complex monitoring information available on all aspects of the system, accessible through REST API or directly visualized in the administration panel 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.



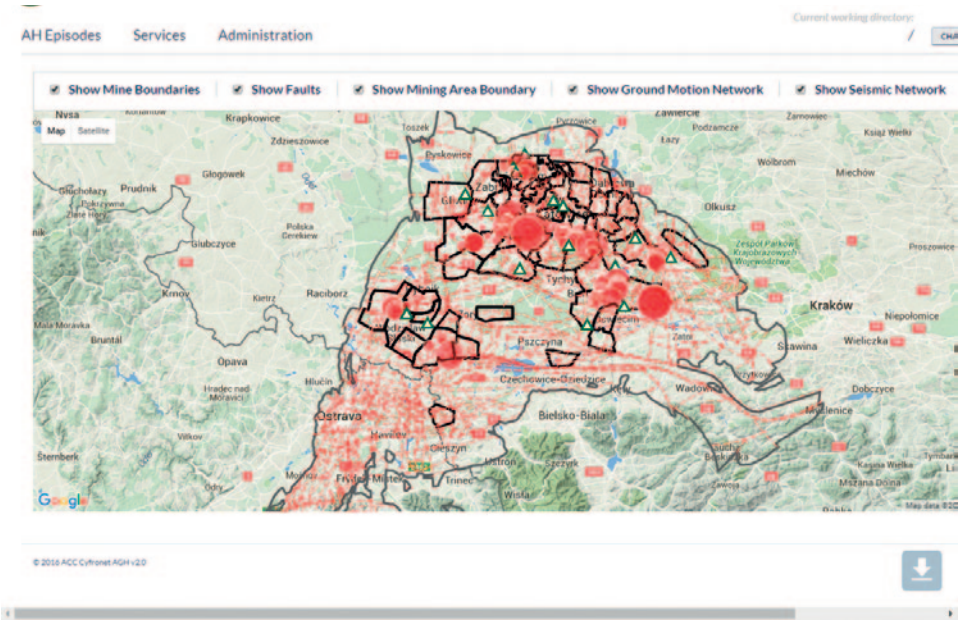
More information: *https://onedata.org*

InSilicoLab is a framework for building application portals, also called Science Gateways. The goal of the framework development is to create gateways that, on the one hand, expose the power of large distributed computing infrastructures to scientists, and, on the other, allow the users to conduct *in silico* experiments in a way that resembles their usual work.

The scientists using such an application portal can treat it as a workspace that organizes their data and allows for complex computations in a manner specific to their domain of science.

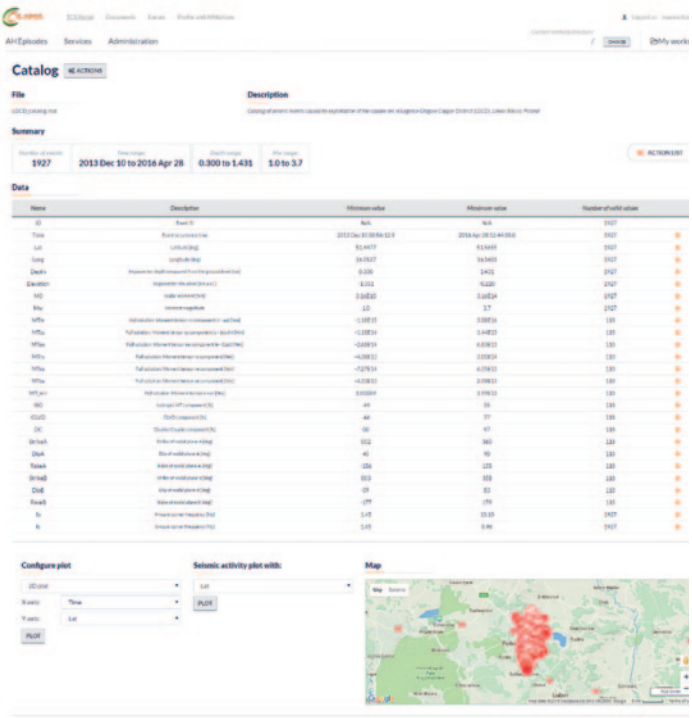
An InSilicoLab-based portal is designed as a workspace that gathers all that a researcher needs for his/her *in silico* experiments. This means:

- capability of organizing data that is a subject or a product of an experiment – this should include:
  - facilitating the process of preparation of input data for computations,
  - possibility of describing and categorizing the input and output data with meaningful metadata,
  - searching and browsing through all the data based on the metadata,
- seamless execution of large-scale, long-lasting data- and computation-intensive experiments.



The approach

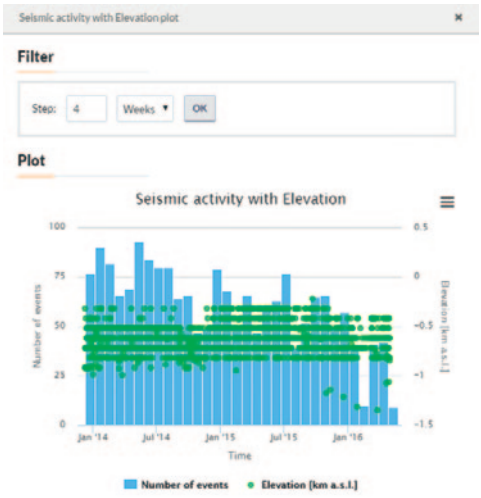
InSilicoLab is not meant to be a "*Yet another engine for...*", therefore, its developers has put maximum stress on the utility of the tool. This means user-friendliness, but, even more importantly, serving real scientific problems. This requires focusing on solving specific problems, rather than building a platform to solve any scientific problem, as the latter cannot be done in a universal and comprehensive way. Therefore, building a framework, which obviously is a generic solution, has to be performed in a bottom-up approach – starting from the particular problems, and building the generic tool from the common parts of the specific solutions.



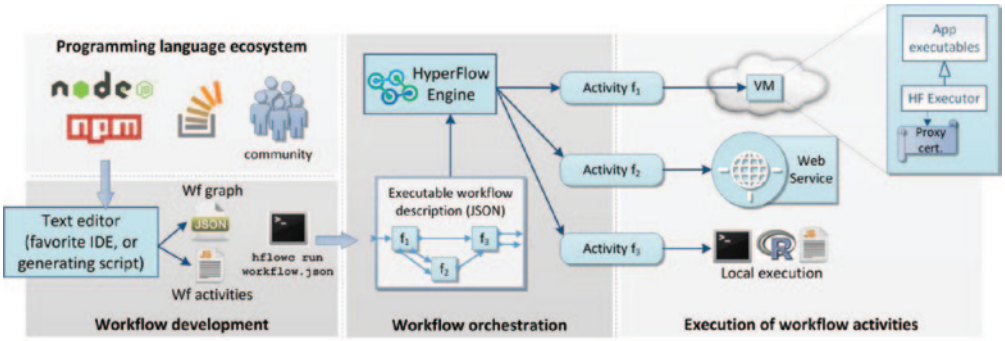
Domain applications

Every gateway based on the InSilicoLab framework is tailored to a specific domain of science, or even to a class of problems in that domain. The core of the framework provides mechanisms for managing the users' data – categorizing it, describing with metadata and tracking its origin – as well as for running computations on distributed computing infrastructures. Every InSilicoLab gateway instance is built based on the core components, but is provided with data models, analysis scenarios and an interface specific to the actual domain it is created for.

http://insilicolab.cyfronet.pl



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.

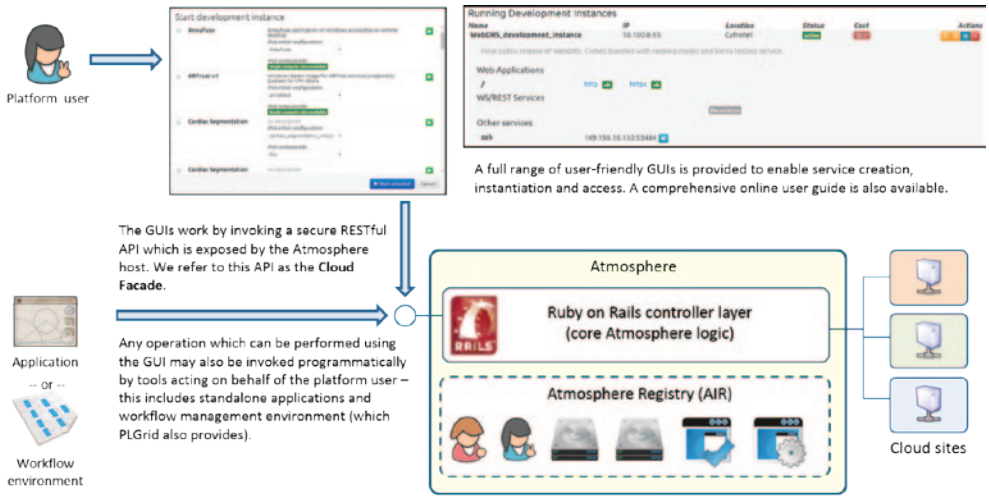
Contact

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*.

Atmosphere

The Atmosphere Cloud Platform is a one-stop site for management and interaction with the computational cloud resources operated by the PLGrid e-infrastructure. Atmosphere is a user-friendly environment where hybrid cloud resources contributed by various participating institutions and sites (including public clouds) are seamlessly integrated into a coherent, unified resource space, made available to PLGrid users. Accessing the Atmosphere platform enables you to exploit the cloud computational resources which are part of PLGrid.



Cloud service abstraction

The principal goal of Atmosphere is to make interaction with cloud sites easy for beginners and experienced users alike. Atmosphere can deploy virtual machines into the cloud, create snapshots and support sharing of computational services by PLGrid groups without the need to use any low-level cloud service libraries. A convenient GUI encapsulates all features offered by Atmosphere to each class of users: system administrators, application developers and end users. In addition, a set of APIs is provided to enable integration of the Atmosphere platform with external applications, tools and workflow management systems.

Security

The Atmosphere cloud platform is integrated with PLGrid authentication and authorization mechanisms. All users of the PLGrid infrastructure can request access to cloud resources simply by joining the **plgg-cloud2** user group. Thereafter, each user can create, use and share cloud-based virtual machines in the context of their own research team, or for individual research purposes. The visibility of each virtual machine (and the corresponding VM images) is restricted to the PLGrid team, in which the machine was created. Atmosphere performs automatic billing and resource consumption auditing for all VMs.

Resources

Atmosphere can interact with many different types of cloud resources contributed by individual cloud platforms. For the purpose of PLGrid, a dedicated cloud site has been set up at ACC Cyfronet AGH, comprising of computational nodes managed by the OpenStack cloud middleware. Atmosphere can also interact with public cloud providers, such as Amazon, RackSpace, Google Compute and many others. All this is done without forcing the user to learn any technology-specific libraries or APIs.

Applications

In addition to raw OS templates for service developers, the PLGrid cloud site supports a variety of ready-to-use applications. Any web or REST service can be deployed into the cloud using Atmosphere. It is worth noting that an earlier version of Atmosphere, developed in the framework of the VPH-Share project was successfully exploited in the Virtual Physiological Human community by approximately 25 research teams affiliated with the VPH-Share and VPH-DARE consortia, as well as external partners who collaborate with ACC Cyfronet AGH. Additionally, Atmosphere was used to provision computational services for the EurValve project in which ACC Cyfronet AGH participated as a member.

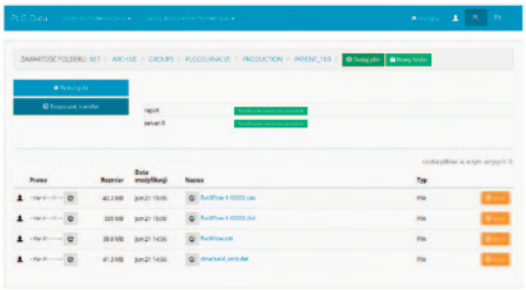
https://cloud.plgrid.pl

Simple tool for file management on a computing cluster

PLG-Data

PLG-Data is a tool for 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 management of access rights for members of a research group, or external collaborators. It is integrated with file systems of both Cyfronet computing clusters: Prometheus and Zeus.

The set of functionalities built in 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). Application of such techniques allows the user to manage their files in a secure way, and, from the point of view of other users of the PLGrid infrastructure, limits access to one's files only to entitled collaborators. In other words, a person who uses PLG-Data does not receive any additional rights of access to files stored in the computing cluster, apart from the rights that the person already has.

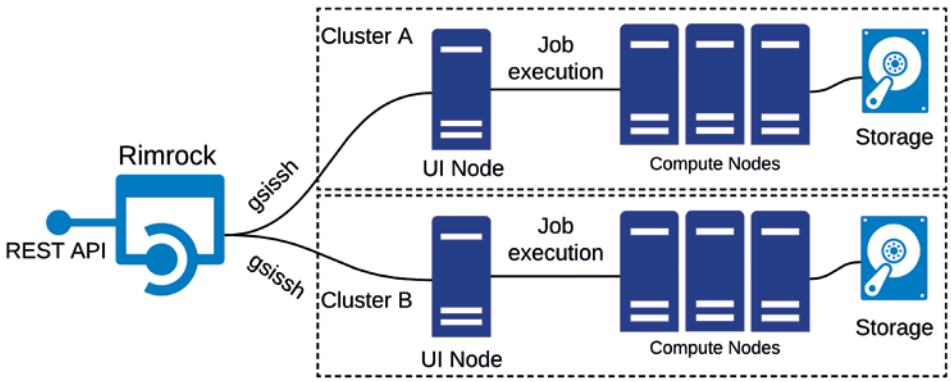
Logging in to the tool is done using either the PLGrid user-password pair, or a p12 certificate installed in one's browser. The service is available either in Polish or English. An advanced programmer's interface (API) helps developers to integrate their platforms, tools and services with the file storage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

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

Rimrock, one of the services of the PLGrid infrastructure, enables 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. By using REST, access to services, applications and advanced scripts deployed on the infrastructure becomes straightforward.

Readiness for various applications

Applying REST principles in the implementation of the rimrock service allows 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, minimizing the role of server-side software.



Support for several job management systems

The rimrock service uses several job management systems (like PBS/TORQUE or Slurm), what ensures support for their 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.

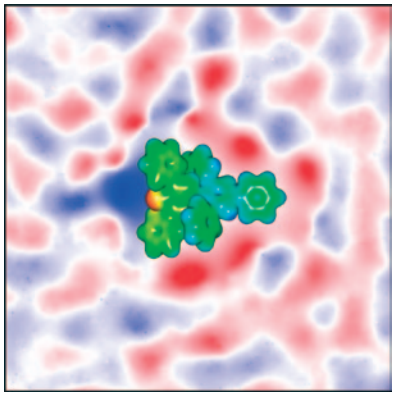
Data security

Data exchanged with the rimrock service is transferred with secure HTTPS connections and for user authorization 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 those being current challenges of nanotechnology or biological sciences) tractable and improve the accuracy of obtained results. Fortunately, constant progress in computer technology together with specialised software offered by Cyfronet meet this demand and enable various kinds of chemical computations.



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

Zeus and Prometheus clusters' nodes provide up to 1.5 TB of RAM and 64 cores per physical node, which enables quantum chemical computations that require large amount of memory or high number of cores with shared memory. Moreover, 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 to storage systems. The parallel-distributed Lustre scratch file system and possibility to use RAMDisk on selected nodes enable that.

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

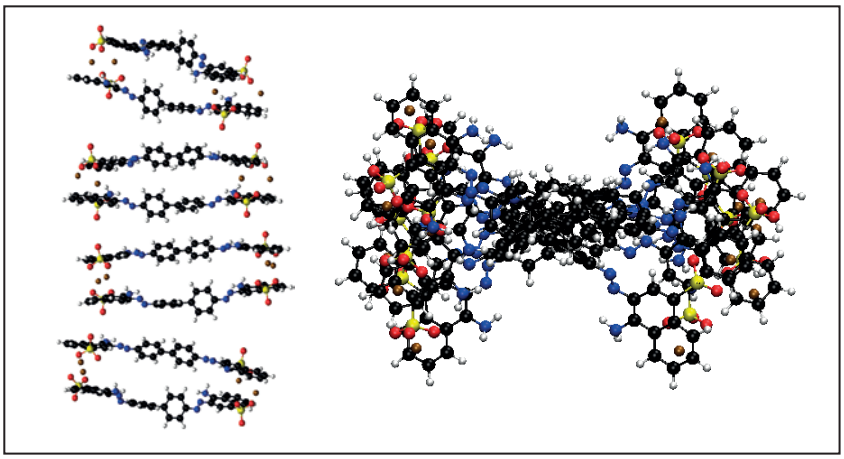
- Versatile and widespread used quantum chemistry codes such as **Gaussian**, **GAMESS UK**, **NWChem**, **Schrödinger**, **Q-Chem** and **TURBOMOLE**, which are capable of calculating electronic structure and various properties of diverse molecular systems using both *ab initio*, density functional 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 **(ADF, DFTB, MOPAC, COSMO-RS)** which 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 (**ADFInput**, **ADFViev**, etc.) ADF package is used by many of our users.

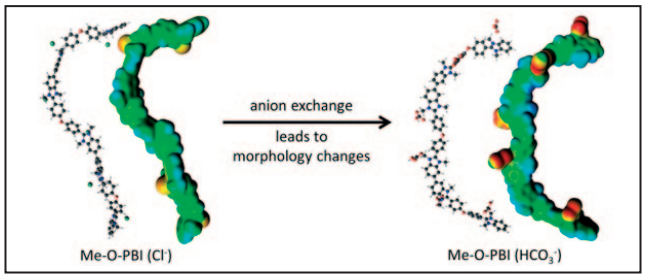
- Several packages, which could be used for **solid-state systems**. Among them **BAND**, **Quantum ESPRESSO** and **SIESTA** are worth mentioning.

- **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. Among software prepared to run on graphical processors our administrators' team prepared quantum chemical packages such as **GAMESS**, **Terachem**, **NAMD**, and **Quantum ESPRESSO**. 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 optimization.

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 optimization 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.

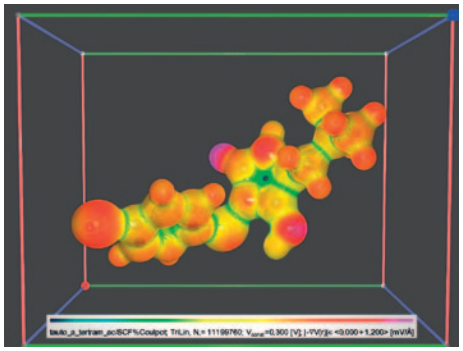
Scikit-learn is a software machine learning library for the Python programming language. It features various classification, regression and clustering algorithms including support vector machines, random forests, gradient boosting, k-means and DBSCAN. It has been designed to interoperate with the Python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

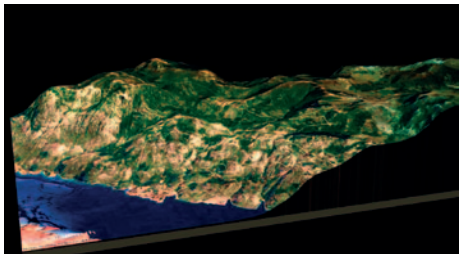
Data visualization enables analysis and understanding of the results of even very complex numerical calculations, especially multidimensional or time-dependent. Most applications for numerical calculations have a module that generates their visualization. Python has a matplotlib or VTK+ module, while Matlab or R also have graphic libraries. The situation is similar with regard to geo-visualization 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 visualization 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 visualizations, 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 specialized 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:

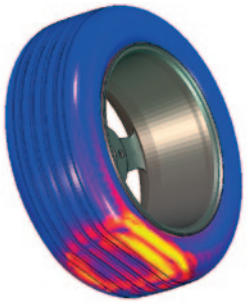


Janusz Opiła: Terrain visualization based on altitude data collected on an irregular grid and textures obtained from the Google Earth Pro application vicinity of Karlobag, Croatia). Own study: DOI: 10.23919/MIPRO.2018.8400037

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

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 nowadays essential tools in the process of developing and building almost everything – from car parts to buildings. Through computer simulations, engineers can check 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 simulations package with 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 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.

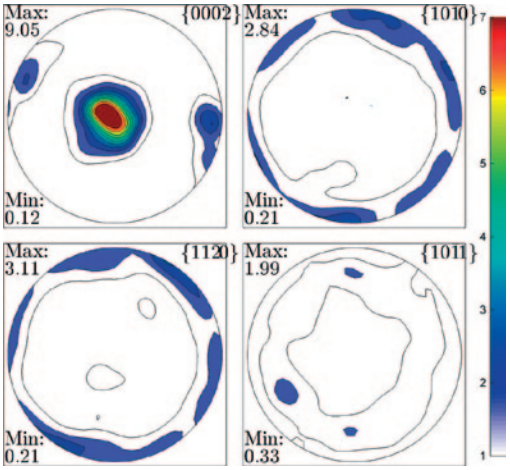
FLUENT software offers the broad physical modeling capabilities needed to model flow, turbulence, heat transfer and reactions for industrial applications ranging from air or liquid flow to semiconductor manufacturing. FLUENT can be used in numerous science domains, including chemistry, metallurgy, biomedicine,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

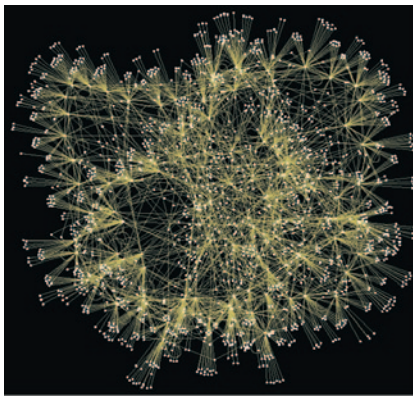


Bartosz Sułkowski: Results of texture simulations by visco-plastic self-consistent model of Zn after hydrostatic extrusion at 250 °C

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 optimization and much more. It is also possible to use Simulink – the environment oriented for simulations and visualizations from blocks, without the need for traditional programming.

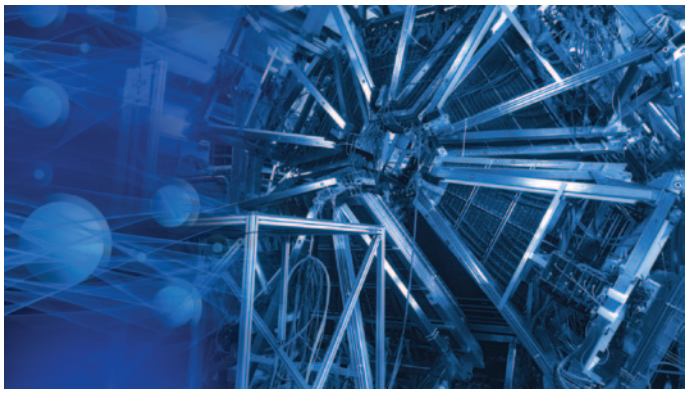
Apart of that environment, users can find in our software a useful application, **MATHEMATICA**, which allows parallel computations with defined precision, dedicated for symbolical and numerical calculations. An advantage of MATHEMATICA is, among other things, a tool for fixing mistakes.

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 team has developed a programming environment, InSilicoLab, which includes a set of advanced tools and programming libraries that allow for the construction and development of dedicated research portals. Such 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 user experiments, even if they are complex, long and requiring many calculations,
- the ability to conveniently describe, categorize and search for input or output.

The InSilicoLab technology is distinguished by striving to 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 laboratory establishes cooperation with renowned scientific and research units as part of Polish and international projects. The effects of cooperation include:

- **Development of the IS-EPOS 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.
- **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.

Contact: Tomasz Szepieniec, t.szepieniec [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 analyzes 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

Laboratory of Visual Techniques (Media Lab 5.0) is a film and television studio whose main task 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 is a film and television studio with an area of approximately 100 sq.m., equipped with a grate with lighting, sound system, with an artificial horizon and scenery. The studio is connected with an auditorium for about 50 people. The studio is also connected with the control rooms of: VISION and LIGHT (with stations of: adjustment of camera tracks, vision realization, recording and playback, computer graphics and light), SOUND with the Yamaha LS 9 console and LECTOR. There are two ASSEMBLY ROOMS with computers equipped with Adobe CS6 software in the Laboratory. The laboratory has Panasonic AG-HPX 371 cameras (3 pcs.) and other film and television equipment for image, light and sound production in studio and outdoor productions.

The studio cooperates with professional creators and producers as well as students of Krakow's universities.

The films made by MediaLab 5.0 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 Acceleration of Computing and Artificial Intelligence

Our mission

Thanks to our expertise in the field of the Artificial Intelligence (AI) algorithms, and our knowledge of the modern computational methods, we can support 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 AI-dedicated partition of Prometheus' supercomputer available at ACC Cyfronet AGH. The AI partition of Prometheus was built based on the four efficient computing servers; each equipped with eight nVidia Volta V100 GPGPU cards. The total computing power of the partition is over 4 PetaFlops, which is over four quadrillions ($4x10^{15}$) of AI-dedicated computations per second.

Domain specializations

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

Natural language processing

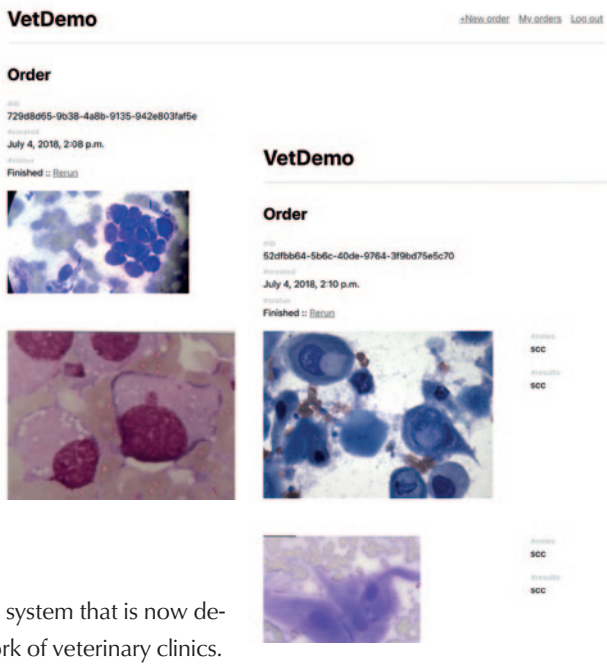
The Laboratory is proud of many years of experience in natural language processing. As an example, we have created and developed the tool that allows the users to search, compare and classify text documents. The result of this work is the web service called Scholar that is available at ACC Cyfronet AGH on the PLGrid platform. One of the important research problems is the analysis of impact of the methods for reducing the accuracy of textual data representation on the effectiveness of the NLP algorithms. We have managed to develop alteration 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 networks-based solutions has revolutionized the NLP field. We research on 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 operation of the system. Our research has shown that it is possible to develop a solution with an accuracy of up to 98.9%.

Image processing

In the field of image processing using neural networks, the Laboratory's work focuses on the recognition and detection of objects for the needs of medical applications. During the tests, we developed

a system for the classification of neoplastic changes in the samples from the cytological examination of the animal tissue. The proposed system achieved an efficiency of approximately 96% for the three selected types of cancer. We used deep network models based on Resnet-50 and Resnet-152 nets. Also, to allow for the selection of the perception area in classification operations, we have developed a special training scheme that is based on genetic algorithms. One of the latest developments in the area is the Yolo3-based detection, which achieves a very good value 0.86 of mAP for the specially prepared images that are made with the use of many low-quality cytological preparations. The system that is now developed is to be ultimately used in the daily work of veterinary clinics.

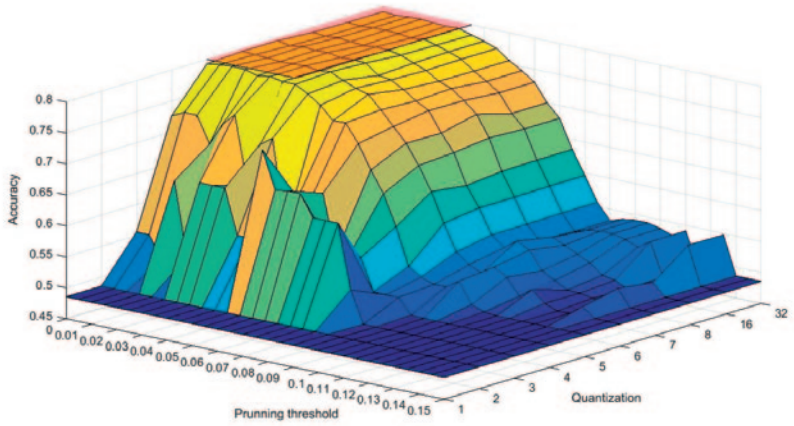


Pictures of cytological preparations showing neoplastic changes of the dog's skin

Time series analysis

The works of the Laboratory also concern the modelling of time series for the detection of unusual situations. Detailed work includes practical applications such as anomaly prediction to avoid catastrophic damage to the magnets and other associated devices of the Large Hadron Collider (LHC) at CERN. By means of GRU and LSTM neural networks and dedicated post-processing, we have developed a system for quench detection in superconducting magnets, which additionally enabled the classification of detected anomalies. Moreover, to allow the user for very low response latency, the developed algorithm was implemented on the computing platform that is based on the Xilinx Zynq UltraScale+ MPSoC 285 FPGA architecture.

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



HPC Users' Conference (KU KDM)

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.

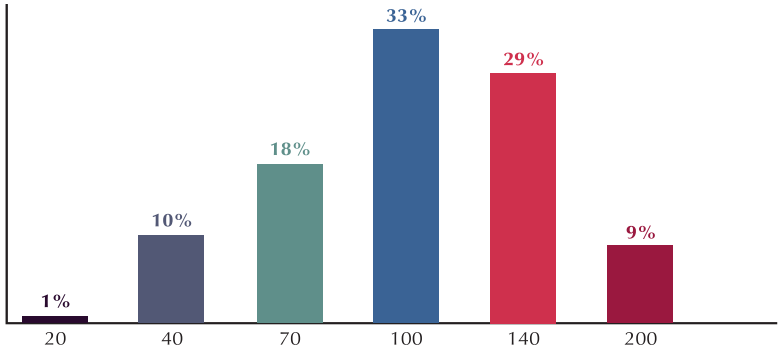
The conference includes a series of talks by scientists who perform research using Cyfronet resources and can present the role of these resources, typical usage scenarios and performance aspects. The event is an important opportunity for Cyfronet representatives to meet with these scientists and acquire 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 domain-specific 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.

In 2020, due to the ongoing coronavirus (COVID-19) pandemic, the conference was held online. More information: *https://www.cyfronet.pl/doack2020/*

Contributed papers elaborated on the basis of the best conference talks were 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*).

Ministry of Education and Science marks of articles published in 2020 by Cyfronet Users in scientific journals



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 – aims 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 is primarily responsible for the development of the station for receiving and processing data from polar satellites (the station was established in Krakow) and plays the role of the coordinator. The joint effort gives 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, is 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

To effectively monitor the processes ongoing below the Earth's surface and their effects perceptible on the surface, we need to use various tools and research methodologies. 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, what 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 better recognize Earth functioning as a complex system, in which on the one hand natural episodes (as volcanic eruptions, floods or earthquakes) have impact on society and economy, and on the other hand environment is being changed by anthropogenic factors. For this reason, effective research requires multifaceted approach. By integration of European research infrastructures into the transnational system, which is delivering data, data products, software and services, EPOS works on enabling access to so far disperse possibilities. The initiative focuses on ten main, linked with each other, Thematic Core Services (TCS):

Image: EPOS - European Plate Observing System

- Seismology
- Near-Fault Observatories
- GNSS Data and Products
- Volcano Observations
- Satellite Data
- Geomagnetic Observations
- Anthropogenic Hazards
- Geological Information and Modelling
- Multi-Scale Laboratories
- Geo-Energy Test Beds for Low Carbon Energy.

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 the Cyfronet's infrastructure. The cooperation is also significantly important in regard of building data visualisation tools, which are managed from the level of web browser with the use of Graphical User Interfaces (GUI). In the long run EPOS aims to connect research communities within one, multidisciplinary platform, to provide effective and safe access tools, as well as to initiate development of new services based on previous achievements.

It's worth underlining that geophysical and geological data as well as 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, as well as anthropogenic threats related to, inter alia, activities of mines. Through a comprehensive analysis of already observed phenomena, one can better prepare for events that have not yet taken place, trying to limit their negative effects as much as possible.

The EPOS prototype 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-UST takes part in subsequent EPOS projects, supporting the initiative both with hardware resources as with the specialists' knowledge and skills. The parallel activities include several areas, one of which is development of digital platform for Thematic Core Service – Anthropogenic Hazards (TCS AH). The platform, accessible at *https://tcs.ah-epos.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 easier analysis of such processes as: flooding of artificial reservoirs, extraction of raw materials, shale gases or groundwater.

In addition to activities related to the sharing and visualization of data, Cyfronet's specialists develop mechanisms of authentication and authorisation (i.a. by adapting the Unity system) and work on scaling the resources for the increasing number of users. Additionally, Cyfronet is sharing a virtual workspace for data management and for running software from various deliverers.

In turn, as part of the EPOS SP project (EPOS Sustainability Phase), Cyfronet is involved in actions aiming to gain new users for the EPOS infrastructure, including industry partners. Parallel work is in progress, which regards maintaining the compliance with central EPOS services and integration of subsequent software packages used in scientific research.

Additionally, by participation in national projects, Cyfronet works for development of the Polish Research Infrastructure EPOS-PL. The latest effort, within the EPOS PL+ project, regards developing the software that supports modelling and conducting research with the use of AI methods. The resulting platform will enable easy use of data integrated in 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.

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. The aim of EOSC is to share in an easy and transparent way not only research data, but also 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 unification of access and authorization 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 not only from the resources themselves, but also from technical support and training. The ACC Cyfronet AGH team played a key role in this regard, becoming the Portal host and developing the Marketplace website, an extensive catalog 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 since 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 behavior in order to create and implement the best user experience practices. At the same time, new functionalities have been implemented.

Since September 2019, the EOSC Synergy project (*https://www.eosc-synergy.eu/*) has been underway to implement EOSC standards for another 9 national e-infrastructures. Cyfronet, in addition to coordinating activities at the national level, supports 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's aim is to upgrade the EOSC ecosystem so that it supports the European research even better, and to convince researchers to use the resources that 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.

Simultaneously with the development of the Portal and accompanying platforms, EOSC is working on creating transparent rules of cooperation, both with service providers and end users. The creation of formative structures is coordinated by the EOSC Secretariat, within which discussion working groups are created. Their goal is to collect requirements and develop contact forms and procedures.

The whole group of the EOSC projects is funded by the European Union's Horizon 2020 program.

EuroHPC – towards exascale computing

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 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.

Reaching exascale is scheduled for 2023. However, another supercomputers of lower performance have been launched lately to build an infrastructure that will allow to scale the technology and software. By the time this publication is composed, EuroHPC has provided 2 pre-exascale supercomputers: LUMI (Finland) and Leonardo (Italy). The third supercomputer of this scale will start operating in Spain. Additionally, 5 peta-scale supercomputers were launched: Discoverer (Bulgaria), Karolina (Czech Republic), MeluXina (Luxemburg), Vega (Slovenia) and Deucalion (Portugal).

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. Its aim is to deliver the modern infrastructure, in particular supercomputing resources, specialized 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 comparing to Europe and the world. The basic element of the infrastructure being built is the computing power provided by supercomputers. Their use is an essential tool to conduct 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 visualization. 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 for 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 optimizing software running in a supercomputing environment.

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

LUMI

The LUMI supercomputer

Poland, represented by ACC Cyfronet AGH and Ministry of Science and Higher Education, has joined the Finnish-led LUMI consortium, which deployed one of the three planned pre-exascale supercomputers. The other consortium members are: Belgium, Czech Republic, Denmark, Estonia, Iceland, Norway, Sweden and Switzerland.

LUMI, which stands for The Large Unified Modern Infrastructure, also means "the snow" in Finnish. And like the snow, it can completely change the landscape – in this case the one of data-driven research. The sum of the unique expertise and experience of building and operating HPC systems, managing Big Data, as well as developing and using the advanced software, shall result in creating an easily-accessible, efficient, safe and powerful HPC environment.

Pre-exascale supercomputer

The LUMI supercomputer has been deployed at the CSC data centre in Kajaani, Finland. Its configuration has been based on 64-core AMD EPYC™ CPUs and AMD Instinct™ GPUs. The aim of LUMI is to provide more than 550 PFlops of theoretical peak performance for European science and economy, which is the equivalent of 1.5 million modern laptops of good configuration. Jointly with over 117 PB of storage and cloud services, it will give researchers the possibility to perform significantly more complicated calculations in shorter time, and faster process bigger sets of data. The expected result is a rapid development of many research directions, including chemistry, biology, nanotechnology, material engineering and other domains crucial for modern industry. The new supercomputer will also contribute to many achievements in areas such as astrophysics, weather prediction, seismic activity and personalised medicine. In the next few years, LUMI is also to become one of the most recognizable research tools in the field of AI – especially *deep learning*.

"We will be able to use the supercomputer to solve the problems that are unsolvable, taking into account the currently existing machines" – says Marek Magryś, Cyfronet's Deputy Director for HPC. – *The flagship example are brain simulations."*

The LUMI supercomputer's size is comparable to a tennis court and it is using 100% hydropowered energy. The heat generated as a side effect of the machine is used to heat nearby buildings.

At the time of full launch, LUMI will be one of the fastest supercomputing systems in the world.

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

The first phase of installation of the LUMI supercomputer. Photo: Juha Torvinen, CSC

Access to LUMI for Polish researchers

Cyfronet coordinates the tasks related to enabling Polish scientists to use the resources offered by LUMI. The aim is to create an efficient and user-friendly work environment, on the basis of good practices. The users of the PLGrid infrastructure may apply for computing grants on the LUMI supercomputer, just as other consortium members.

Competition for pilot projects

In 2021, it is planned to launch pilot computational grants to test the capabilities of the LUMI infrastructure. Poland, as a member of the consortium, submitted projects of potentially significant importance, requiring the use of a large number of computing cores and massive storage resources. Pilot projects are selected through competitions and will be implemented in two stages:

1. CPU-based projects (data intensive computing, high-throughput computing and high-performance data analysis workloads) – autumn 2021 – 2 projects from Poland (*the competition has been resolved*).
2. Highly-scalable GPU applications – beginning of 2022 – 2 projects from Poland.

Access is provided by the PLGrid Portal (*https://portal.plgrid.pl*) and additional information is available at: *https://cyfronet.pl/lumi*.

Cyfronet in 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.*

As part of the operational phase of the EPOS infrastructure, the resources of the thematic cores are made available. Poland provides a thematic core of anthropogenic hazard, the so-called TCS AH (Thematic Core Services Anthropogenic Hazard). The main task of Cyfronet in the project is the operational provision of the portal and TCS AH resources integrated with it (data, software and computing power for its analysis).

The main goal of the EuroCC project is to create a European technological support system based on closely related National Competence Centers in individual European countries, which will allow the academic environment, the enterprise sector (especially SMEs) and public administration to benefit from the available expertise, experience and resources of EuroHPC.

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.

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

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 EOSC-Synergy project introduces EOSC standards for national infrastructures in nine European Union countries. This will be done by harmonizing policies and expanding access to research infrastructures, scientific data and domain services.

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

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

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 solutions developed within the project are a breakthrough step in the creation of innovative, exascale data processing services, maximizing the benefits of modern data processing systems.

The aim of the EGI-Engage project was to accelerate the implementation of the Open Science Commons by expanding the capabilities of a European backbone of federated services for computing, storage, data, communication, knowledge and expertise, complementing community-specific capabilities.

The project aim was to develop and implement new tools and services used to run interactive applications, which required high computing power and large data collections in the grid environment.

The goal of the project is 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.

The aim of the project is 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 is the Polish stage of development of the EuroHPC program.

Within the project the construction of unique research laboratories based on the national PIONIER fiber optic network is planned. The main goal of the project is 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 is 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 will be increased. A new Research Infrastructure Center (Center for Research Infrastructure of Satellite Data - CIBDS) will be established, a new test site (Geophysical Safety System for mining protection pillars) will be created, and the WNiP will be established: „IT Platform for Research with Artificial Intelligence Methods (EPOS-AI)".

The goal of the project is 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, their secure storage and sharing for the purposes of science, administration and training.

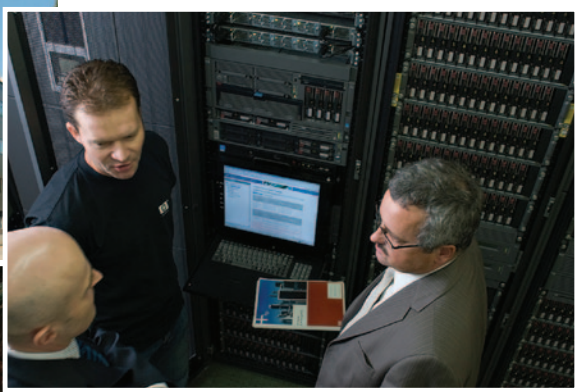
The project aimed at building the national research infrastructure for solid Earth Science and its integration with international databases and services implemented under the European Plate Observing System (EPOS).

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 aim of the project was integration of selected services available in the PIONIER network, and the development of the new services, e.g., with the increased reliability and security.

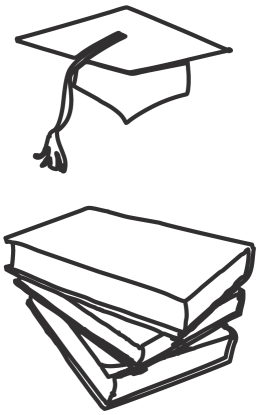
The aim of the project was the development of 21 environmental science data communication networks providing the scientific institutions across the country with access to a modern and secure network infrastructure, supporting the research and development of Polish groups of scientists.

The project envisaged the creation and launch of five services running on the basis of the PIONIER network. These services included: video conferencing services, eduroam services, campus services, universal archiving services and scientific interactive HDTV services.




www.CYFRONET.krak

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. The authors of the best works receive valuable prizes. 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 variety of scientific problems in chemistry, energy sector, electronics, physics, computer science, materials engineering, mathematics and robotics. 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 2022 Contest edition!

http://www.cyfronet.pl/konkurs

The laureates of one of the previous Contest editions

Kévin Almeida Cheminant

The interview with the author of the PhD thesis:
„Search for ultra-high energy photons through preshower effect with gamma-ray telescopes"

What made you devote your PhD dissertation to the issue of gamma photons with extremely high energies?

Ultra-high energy cosmic rays (UHECRs) are charged particles traveling through the Universe with the energy comparable to the energy of a professionally-served tennis ball (Exa-electronvolt - EeV). Such energy is well beyond the reach of any human-made accelerator and therefore, UHECRs are very valuable to understand physics at the most extreme energies. Yet, their origin as well as the mechanisms through which they obtain such tremendous energies are not yet fully understood. However, UHE photons are often emitted in association with UHECRs. Searching for these photons has several advantages. As neutral particles, they are not deflected by magnetic fields and retrieving their source of emission would shed some light on the origin of UHECRs. Secondly, different scenarios for UHECR production lead to different expectations regarding the UHE photon flux. Unfortunately, no UHE photon has ever been observed. My PhD thesis was dedicated to develop a new method to search for UHE photons by using gamma-ray telescopes, which are normally dedicated to the observation of GeV-TeV photons. We showed that these telescopes could efficiently detect UHE photons, allowing the handful of gamma-ray telescopes spread out across the world to join the search already initiated by the Pierre Auger Observatory and Telescope Array.

You are a long-term participant of the CREDO project. How would you introduce it to a person not related to the field?

CREDO stands for Cosmic-Ray Extremely Distributed Observatory. The idea behind this project is to create a worldwide network of cosmic-ray detectors in order to search for possible correlation between distant cosmic-ray detections. In order to create such a network, every potential cosmic-ray detector must be considered, the professional ones (neutrino, dark matter, gamma-ray experiments, etc...) as well as the educational ones (CosmicWatch, higwh-school arrays, etc...). For that purpose, the CREDO collaboration developed a smartphone application capable of turning the camera sensors into a particle detector. CREDO's main philosophy is to make all the data and analyses fully accessible to anyone willing to join or contribute to this project, and to not leave any area unexplored by considering potential connection between cosmic rays and other fields of physics such as biophysics (impact of radiation on living organism) or geophysics (relation between cosmic rays and earthquakes). More information about CREDO can be found on the following website: *https://credo.science*.

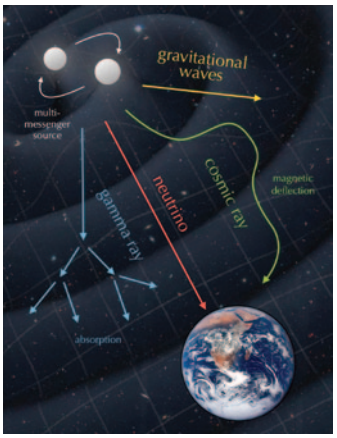
In your dissertation you use different types of data: from observations, and obtained as a result of computer simulations as well. How do you assess the advancement and usefulness of computer modeling in astrophysics, how do Cyfronet's resources match the expectations of astrophysicists?

Cosmic rays and gamma rays interacting with the Earth's atmosphere produce cascades containing hundreds of thousands of secondary particles which are the result of many different types of interactions. As analytical models fail to give a full description of these cascades, we must rely on extensive Monte-Carlo simulations which require large computer resources. These simulations can describe the behavior of each particle and also take into account the probabilistic nature of particle interactions. Thanks to them, we can build expectations and test different physical models, which are then compared to the collected data. In my personal work, Cyfronet's resources have allowed me to run a very large number of extremely demanding Monte-Carlo simula-

tions. Without them, it would have been impossible for me to obtain enough statistics to build my expectations in a reasonable amount of time. Moreover, as our understanding of physical interactions increases and as we aim for more detailed descriptions of physical phenomena, supercomputing facilities will have an even greater role to play in the coming years and I am confident that Cyfronet's expertise will be extremely valuable for any scientist in need of large computer resources.

Based on your experience and knowledge as well as the results of the dissertation, what are your predictions regarding the development of the discipline?

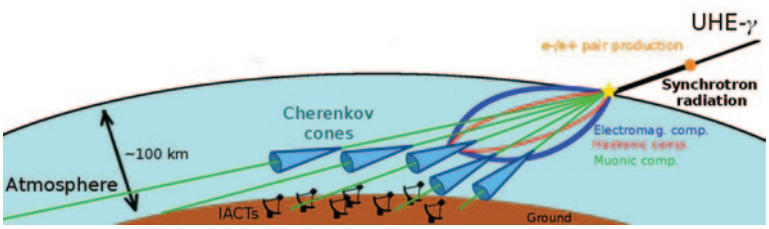
Although existing astroparticle physics experiments have produced an unprecedented amount of knowledge of the Universe at the highest energies, we are still striving for more sensitive observatories with better angular and energy resolution. The next-generation of gamma-ray telescopes proposed by the Cherenkov Telescope Array (CTA) collaboration will provide a better understanding of known gamma-ray sources, but also allow the search for extreme particle accelerators. In ultra-high energy cosmic ray physics, additional detectors of different types at the Pierre Auger Observatory and an increase of the effective area of Telescope Array will provide more information about particle interactions in air showers and larger statistics needed to understand the nature of ultra-high energy cosmic rays. Neutrino experiments such as IceCUBE, KM3NeT or Baikal-GVD also open a new window on the Universe of extreme energies. In fact, the future of the discipline inevitably lies in a multi-messenger strategy that has already been implemented by the astroparticle physics community. Thanks to a system of alerts, flaring astrophysical sources are now observed by different kinds of experiments, which allows us to gain very important knowledge about the physics at stake in these sources.



Astrophysical objects may be sending various types of information, including neutrinos, gravitational waves, cosmic rays, and gamma rays. Observing these objects with different kinds of experiments to detect these messengers provides us with a better understanding of the physical phenomena occurring at the highest energies.

Is there any advice you would like to give the people considering just starting doctoral studies?

I think the best advice I can give to anyone starting doctoral studies is to not be afraid to explore new methods or new topics. Being a novice in specialized fields can be frightening. In my case, I often felt like anything I would come up with would simply be not interesting. I think any research project requires a bit of creativity and curiosity, and one should not fear to experiment his or her own ideas, even if it eventually fails. Asking questions and constantly staying up-to-date with the field are also what can make PhD studies a very successful and rewarding experience. Finally, giving oral presentations is an excellent exercise. It's a good way to summarize efficiently one's own research, but also to gain valuable remarks and comments from experts.



An ultra-high energy photon interacting with the geomagnetic field produces an electron/positron pair which emits secondary photons (synchrotron radiation). This collection of particles, also called preshower, produces an atmospheric cascade containing hundreds of thousands of particles. Such a cascade can be detected on the ground by Imaging Atmospheric Cherenkov Telescopes (IACTs) through the blueish Cherenkov radiation emitted by charged particles (electrons, muons, etc.) in the air.

Klemens Noga

The interview with the author of the PhD thesis:
„Study of electron interactions in molybdenum compounds on mixed valence by methods of quantum chemistry"

What aspect of computational chemistry is most interesting to you?

In my opinion, the most interesting aspect of computational chemistry is that research is carried out on various models that are relatively easy to modify. Thanks to this, we can analyze the influence of various factors on the properties of chemical systems. Additionally, with the advancement of computing technology, more and more complex and sophisticated models could be tackled.

Why did you decide to study scorpionate molybdenum complexes?

Initially, my interest in scorpionates arose from the need of the moment. My collaborators from the Cracow University of Technology worked on understanding the mechanisms governing the electrochemical properties of scorpionate molybdenum complexes, especially those containing more than one metallic centre. They wanted to get an insight into the factors influencing the electronic structure of investigated compounds. Fortunately, I have been able to apply successfully computational chemistry methods to study their complexes. Molecular modeling provided an opportunity to gain additional knowledge about electronic factors that suppress electronic transfer between centers despite their geometrical proximity.

The scorpionates themselves turned out to be a very engaging and complicated subject of research. One aspect is the versatility of building blocks that can form them – diverse ligands and metal centres. Another is the variety of multicentre complexes connected by various bridging ligands. Furthermore, their ability to create compounds with mixed-valence species, where individual metal centres are on different formal oxidation states and unpaired electron may be delocalized between metal centres, makes them fitting systems for investigation of electron transfer.

Molecular modeling has also given insight into the mechanism of electrocatalytic reduction of chloroform by molybdenum scorpionates with alkoxy ligands and its inhibition by alcohols and alkenes investigated by the experimentalist. The role of dispersion interactions in the stabilization of weak adducts of scorpionate complexes with small organic molecules crucial to investigated electrocatalysis has been analyzed.

During the research, an efficient computational protocol for redox potential prediction for investigated compounds has been proposed and tested. The proposed procedure can be used to accurately predict redox potentials of the vast family of molybdenum and tungsten scorpionates and other compounds.

What opportunities did the cooperation with the experimental team give you during theoretical research?

Close cooperation with the electrochemists – the team of Piotr Romańczyk and Stefan Kurek from the Cracow University of Technology – was crucial in my case. Thanks to this, we were able to complement each other›s lack of understanding of various aspects of the problems being solved, and I was able to directly confront the results of my computational experiments with experimental data.

During your research, you have used several subsequent Cyfronet supercomputers. How have those resources influenced your research?

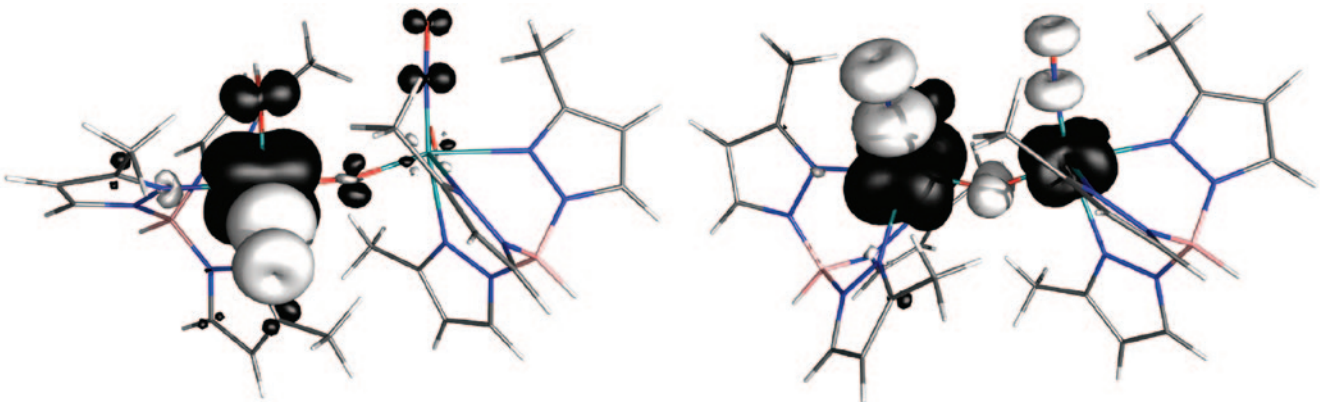
To put it simply – without the resources of Cyfronet, I would not be able to carry out my research. Each successive generation of supercomputers provided me with new computational possibilities. I have been able to study more complex systems or perform computations using methods that allowed me to obtain more accurate results which could better reproduce available experimental data.

Based on the process of your research, could you make an attempt to predict the directions in which modern technological achievements and theoretical discoveries are leading science?

In recent years, the increased importance of *in silico* experiments in chemistry, biochemistry, and related sciences can be noticed. Thanks to the use of supercomputers, interesting systems can be investigated faster and more effectively. Also, so-called machine learning has contributed to obtaining interesting scientific results in biochemical sciences, enabling, for example, the prediction of spatial structures of proteins. I think we will be able to create better, more complex models, which will give us a better insight into the most important chemical and biochemical processes.

The research leading to a dissertation is spread out over time and ends with a different set of knowledge and skills than the one with which it began. What advice would you give young scientists at the beginning of this path?

I would recommend that they devote the first period of their studies to expanding their knowledge about available research methods and techniques and establish the most comprehensive contacts with various scientists. Thanks to this, already at the stage of doctorate, one can participate in very compelling research projects, see the methodology of many research groups, and have the opportunity to observe and participate in the preparation of various reports and publications. Only then, based on my previous experience, I would choose a specific thesis subject on which I would work in the following years. Additionally, I would advise anyone starting such an adventure to focus also on the art of scientific writing. This skill will be beneficial not only when writing a doctoral dissertation.



The contour of the spin density computed for one of the investigated molybdenum scorpionates in the geometry of the energy minimum (left) and the maximum delocalization of the unpaired electron (right)

Tomasz Prokop

The interview with the author of the PhD thesis:
„Three-dimensional numerical analysis of transport phenomena in a Positive-Electrolyte-Negative assembly of a Solid Oxide Fuel Cell"

What made you devote your doctoral dissertation to the issue of obtaining electricity from SOFC cells?

My interest in Solid Oxide Fuel Cells (SOFCs) was sparked by my research supervisors, back when I was a Master's student at our department. SOFC research is full of interesting, multidisciplinary research challenges. Besides, the use of SOFCs is one of the most efficient ways to utilize the energy of fossil fuels. Solid Oxide Fuel Cells have exceptionally high efficiency of converting the chemical energy to electrical energy, and SOFC based devices may be operated with a wide variety of fuels, including natural gas. The existing 1-250 kW SOFC systems, which can be used as auxiliary power sources of residential buildings, office buildings or server plants, can match the efficiency of ~1 gigawatt blocks of large scale gas power plants. At the same time, they are silent, they do not emit poisonous substances, and their waste heat may be easily used for heating. Proposed SOFC power plants would use roughly two-thirds of the fuel used by an equivalent conventional installation (and emit an appropriately lower amount of pollutants).

How can research be translated into practical implementations?

My research is focused on modeling the phenomena occurring in the microscopic structure of Solid Oxide Fuel Cells. The code I wrote may be helpful during the design process of SOFC devices. My program uses digital representations of SOFC microstructures (which can be obtained by investigating a real electrode, or generated artificially by another program). The program analyses voltage losses related to electric resistance, gas transport inefficiency and other factors. Such a simulation may be used to predict how a change in cell's features (such as material parameters, electrode composition, layer thickness, and others) will affect the device performance at various conditions, without the need to manufacture a prototype and conduct a series of arduous experiments. I believe that my research will help to increase the efficiency of SOFC cells, while decreasing the cost of their development, eventually resulting in better rate of return for commercial SOFC systems.

What are the next challenges for you and other scientists studying fuel cells?

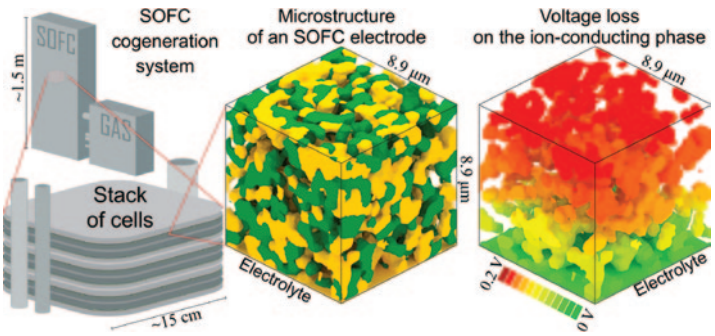
The major challenges of SOFC industry include developing mass production and quality assurance techniques which would allow for lower cost and higher reliability of commercial SOFC systems. Since SOFCs function in temperatures as high as 1100°C, it is necessary to use heat resistant materials and sealing with good resistance to corrosion. Solid Oxide Fuel Cells are still somewhat fragile, and the thermal stresses, which are related to heating and cooling, limit the number of cold shut-downs during the device's lifetime. Furthermore, the microstructure of fuel cell electrodes evolves during long term operation, which affects their performance significantly. This process is yet to be fully understood. Research problems related to fuel purifying, reforming, storage and delivery are also very important for wider application of fuel cells.

Your doctoral dissertation is international, literally intercontinental. Can you recommend this approach to colleagues who are starting their adventure with doctoral studies?

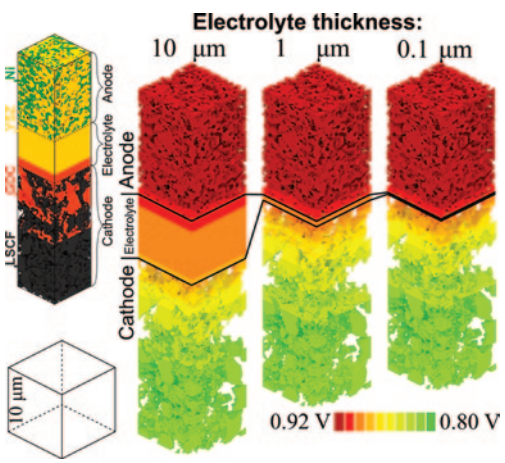
Absolutely. Cooperation among researchers is inevitable in science, and a compatible research team may be located in another country. It is the norm everywhere, excluding only the largest countries with a significant R&D sector. A typical scientist is focused on a relatively narrow research problem – that particular problem is often taken up by only a few research teams in the entire world. It is helpful to contact at least some of those researchers. They might have overcome research challenges which are particularly troublesome for us, and our findings may be useful to them as well. International cooperation improves the access to equipment, as well as internships, and scholarships opportunities. It also increases the visibility of publications, and helps to determine novel, interesting research directions.

Finally, could you briefly characterize the Japanese approach to university education and research? Any striking cultural differences?

Unlike Polish Master's students, Japanese Master's students normally do not participate in many courses. Their studies resemble the work of Polish PhD students. They are assigned to a specific research team, headed by a professor. They have their own desks and unrestricted access to laboratory equipment. One of the most important aspects of my internship in Tokyo was the participation in weekly staff seminars. During these meetings each employee, each PhD student and each Master's student delivered a presentation detailing his recent research work or discussing interesting papers published by other teams. These meetings took several hours, but they were very interesting. Team building events are also very common, e.g. weekend-long tours. Welcome parties are organized for new team members and farewell parties are held for those who are leaving. The approach to cleaning is different, too. The cleaning staff of the University takes care of the common areas, however laboratories and offices are kept tidy by the researchers and students themselves. Many Westerners find it peculiar that it is socially acceptable for the Japanese to take naps at work. Culturally, it is viewed as a sign of tiredness due to diligent work. This custom is particularly understandable at research institutions – which are open 24 hours a day. Some researchers can be seen working even in the middle of the night.



The microstructure of a SOFC electrode in context of an example device. Every phase in the microstructure has a specific function – voltage loss on the ion-conducting phase shown



A parametric study: ion-conducting phases potential for different values of electrolyte thickness

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

Cyfronet's main mission is to serve the scientific community and to actively support also individual researchers, by providing them with the IT resources of the Centre: starting from the computing power of supercomputers, through mass storage, to a super-fast computer network and useful IT tools. We are very pleased and proud that among the wide range of topics and research work carried out with the help of ACC Cyfronet AGH supercomputers, some have been appreciated and awarded by the Polish National Science Centre. The scientific successes of Cyfronet infrastructure users are the best proof of its usefulness as a tool for conducting scientific and research work.

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 Michał Tomza, PhD DSc from the Faculty of Physics of the University of Warsaw*

On October 14, 2020, Michał Tomza won the Prize in the field of technical sciences. The prize was awarded for the theoretical description of interactions and collisions between ultracold atoms, ions and molecules.

Michał Tomza is a physicist and theoretical chemist, assistant professor at the Institute of Theoretical Physics, Faculty of Physics, University of Warsaw. He specializes in the quantum description of matter at ultra-low temperatures, including the theory of interactions and collisions of ultracold atoms, ions and molecules. As part of research internships, he conducted research at the University of Granada, the University of British Columbia in Vancouver, Harvard University, the University of Colorado in Boulder, the University of California in Santa Barbara and the University of Innsbruck. In 2017,

he created his own research group at the University of Warsaw, where he received research grants in the SONATA and OPUS competitions of the Polish National Science Center as well as Homing and First Team from the Foundation for Polish Science.

Dr. Tomza is the research supervisor of several graduate students and doctoral students. He works closely with experimental groups from Germany, the Netherlands and Switzerland. Recently, his research interests have focused on ultracold mixtures of ions with atoms.

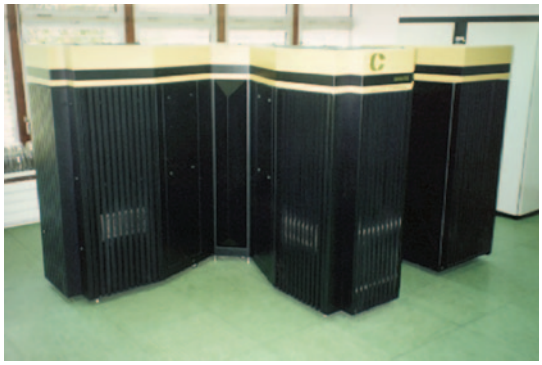
– *In my scientific work I study the properties, dynamics and potential applications of matter at ultra-low temperatures* – says Michał Tomza. – *Ultracold matter consists of atoms, ions or molecules with a temperature close to absolute zero, i.e. much lower than, for example, the temperature of outer space. In such conditions, the quantum nature of the world is clearly revealed, which is why my research allows me to gain new insight into the theory of matter and the interactions between matter and light – topics important both from a cognitive and technological point of view. At the microscale, at the atomic level, the entire world around us is described by the laws of quantum physics, and a number of devices in use today, such as a laser, tomograph or GPS, could not exist without quantum effects. We are now also on the verge of the birth of new quantum technologies that will directly use the fundamental laws of nature to guarantee maximum detectors' sensitivity, maximum computers' performance, and maximum encryption security. My goal is to develop concepts and theoretical methods for quantum description and discovering the properties of matter at ultra-low temperatures. The theoretical results I get are used in new experiments at the border of physics and chemistry, and contribute to the development of new quantum technologies.*

Michał Tomza supports his research with calculations on the Prometheus supercomputer at Cyfronet, within the *plgquantmol4* computational grant: "Simulations of ultracold quantum multi-body molecular systems".

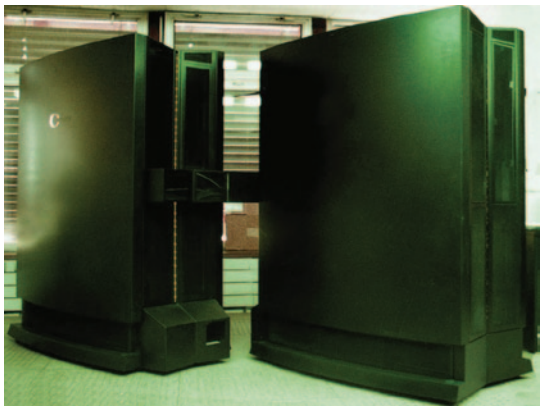
Congratulations to the winner and we wish you further success!

**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 KraKow 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 Krakow and Warsaw.

Construction begins in the Krakow MAN

1994 A 2 Mbps link is deployed between Krakow 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 Krakow 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

SGI Origin2000

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

SGI ALTIX 3700

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 81st position on the TOP500 list

HP Cluster Platform 3000 BL

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

In April, ScaleMP, a leading provider of virtualization 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 Rzeszow through the PIONIER network link of 2x10 Gbps capacity

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, 2.65 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 3.5 PFlops is deployed in Cyfronet.

Prometheus and Ares rank 373rd and 216th places on the TOP500 list (the June edition).

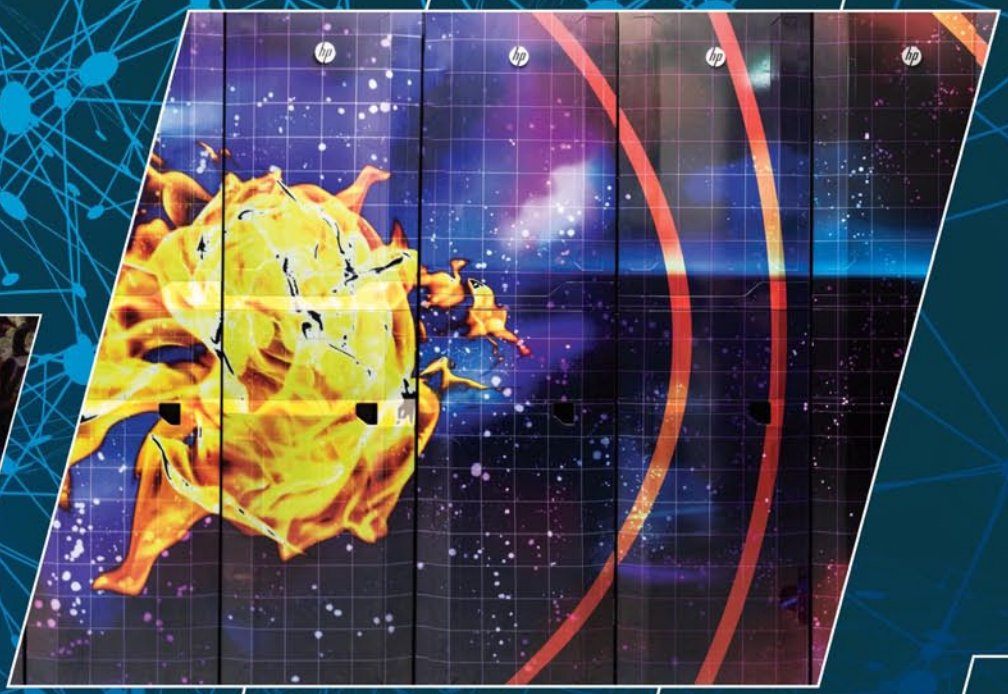
A new version of the ACC Cyfronet AGH website has been launched.

EuroHPC PL, PIONIER-LAB, KMD3, EGI ACE and EOSC Future projects launched.

AGH UST CAMPUS

LEGEND

1. Rector's Office
2. Faculty of Mining and Geoengineering
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. Walery Goetel School of Environmental Protection and Engineering
22. Department of Foreign Languages
23. Department of Sport and Physical Education
24. AGH UST Swimming Pool
25. Centre of e-Learning
26. AGH UST Academic Computer Centre CYFRONET AGH
27. University Computer Centre
28. Department of Education
29. Centre for International Students
30. Education Centre
31. University Admissions Board for Prospective Students
32. AGH UST Student Campus
33. University Board of Student Government
34. Career Centre
35. Centre for Transfer of Technologies
36. Administration and Business Cooperation Department
37. Krakow Centre for Innovative Technology INNOAGH
38. Centre for Project Management
39. Department of International Collaboration
40. Disability Support Office
41. AGH UST Museum
42. Geological Museum of the Faculty of Geology, Geophysics and Environmental Protection
43. AGH UST Press
44. Academic Cultural Centre, Club STUDIO
45. Student Club Gwarek
46. Student Club Zaścianek
47. Student Club Karlik
48. Student Club Filutek



ISBN 978-83-61433-39-2

Academic Computer Centre
CYFRONET AGH
Nawojki 11

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

cyfronet@cyfronet.pl
www.cyfronet.pl