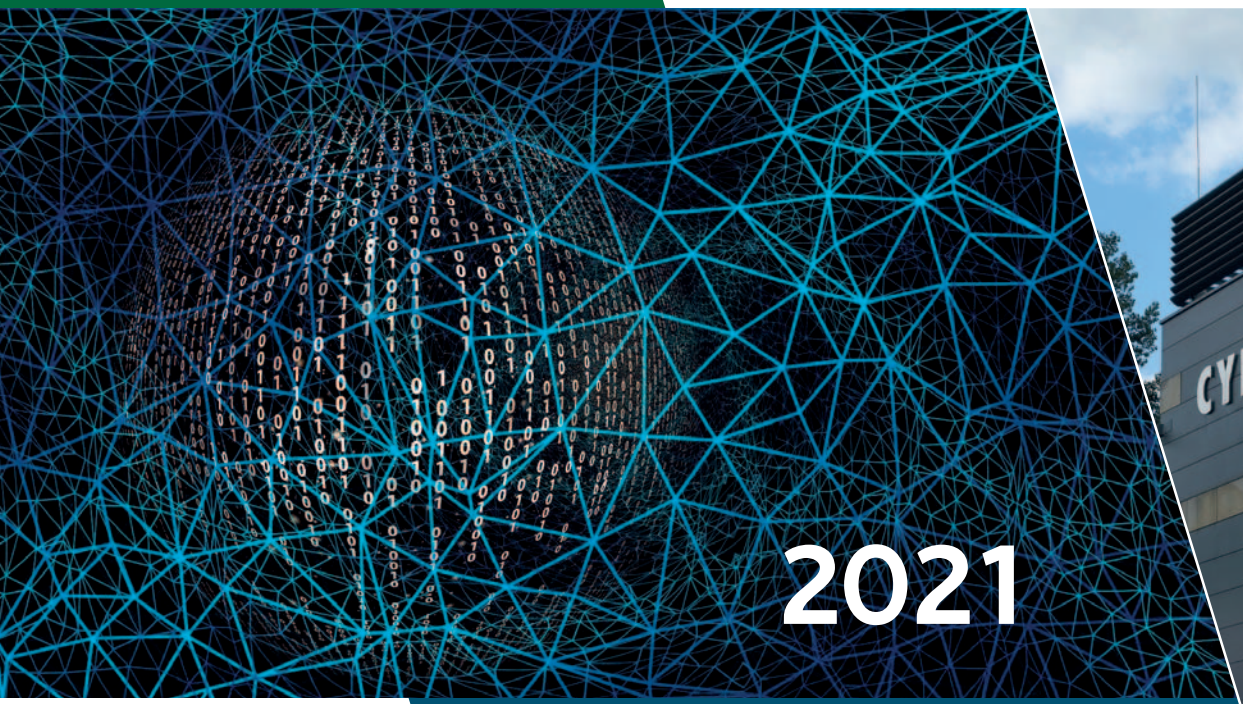
Academic Computer Centre

CYFRONET AGH



2021

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!

The past year brought a new challenge to Poland and the world in the form of the COVID-19 pandemic. The response to this situation requires intensified cooperation of specialists from many fields. At ACC Cyfronet AGH, operating in the areas of information technology, science and economy, we constantly support scientists in their research, also in these difficult times. Moreover, computational grants dedicated to research on the SARS-CoV-2 coronavirus and the COVID-19 disease it causes have priority in the access queue for Prometheus resources and are given the highest priority in running. Thousands of supercomputer computing cores and accompanying infrastructure allow for fast processing of large amounts of medical, biological and chemical data. The carried out studies concern antibodies present at the time of infection, molecules with potential for inhibiting infection, and opportunities for vaccine development.

In June 2020, the Prometheus supercomputer with 2.65 PetaFlops of computing power, again as the only supercomputer from Poland, took the high 288th place in the TOP500 list of the world's fastest supercomputers. In addition, after several years of launch, it still remains one of the most energy-efficient machines, which was confirmed by taking 142nd place on the Green500 list. In 2019, Prometheus and its predecessor Zeus altogether performed almost 5 million computing tasks for scientists representing many universities and research institutes.

The Prometheus supercomputer has been used in medicine for a long time – specialized software packages installed in its resources are used on a daily basis, for example in research related to modeling of drug molecules, for analyzing data from DNA microarrays, identifying proteins and predicting their role in biological processes.

Cyfronet has been supporting initiatives related to the development of medicine and pharmacy for many years. It is part and is the initiator of the consortium implementing the SANO project, one of the three projects submitted by Polish research units, which were awarded grants in the European Teaming for Excellence competition. The aim of the project is to introduce completely new diagnostic and therapeutic solutions resulting from the production of new, computational biomarkers of diseases, individualized for particular patients.

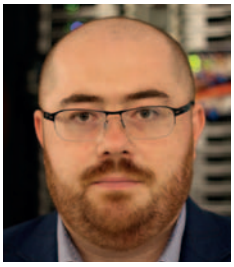
However, we cannot forget that scientific research carried out with the help of the Centre's resources concern many fields, and the number of scientists using our supercomputers is constantly growing. We are glad to support researchers in their scientific work and in making important discoveries. Each success of Cyfronet infrastructure's Users is extremely important to us. In turn, what may be important from the Users' point of view is the fact that 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*.

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 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 Storage Department – Marek Magryś, MSc (manager)
Phone: +48 12 633 34 26, e-mail: Marek.Magrys@cyfronet.pl

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

High Performance Computing Users Department – Mariusz Sterzel, PhD (manager)
Phone: +48 12 632 33 55, e-mail: Mariusz.Sterzel@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 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 Visual Techniques – Jacek Przybylski, MSc (manager)
Phone: +48 12 632 33 55, e-mail: Jacek.Przybylski@cyfronet.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

photo: Adam Frączek, KSAF AGH

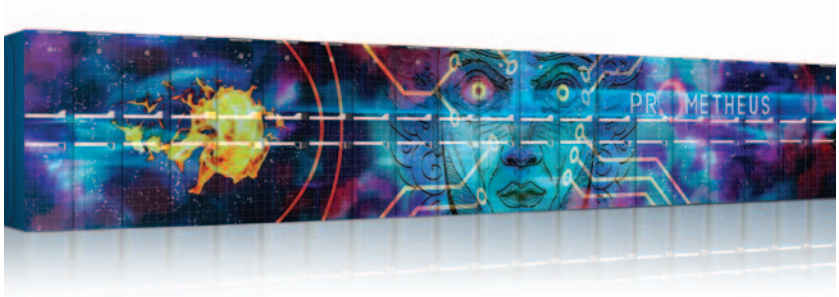

photo: Krzysztof Mastalski, KSAF AGH

Term of office: 2017-2020

Marek Jarnicki, Prof. *The Chairman*	Jagiellonian University
Marek Jeżabek, Prof. *Deputy Chairman*	Institute of Nuclear Physics Polish Academy of Sciences
Bogdan Bochenek, MSc	Institute of Meteorology and Water Management
Andrzej Dietrich, MSc	Institute of Petroleum and Natural Gas
Jan Dudek, MSc Eng.	Cracow University of Economics
Florian Gambuś, Prof.	Agricultural University in Cracow
Jerzy R. Jaworowski, PhD Eng.	Cracow University of Technology
Paweł Kruszelnicki, MSc Eng.	Academy of Physical Education
Robert Gryboś, PhD	Jerzy Haber Institute of Catalysis and Surface Chemistry Polish Academy of Sciences
Wojciech Majka, MSc Eng.	Jagiellonian University Medical College
Paweł Mikłaszewski, MSc	Academy of Music
Roman Mazur SDB, PhD	Pontifical University of John Paul II in Cracow
Andrzej R. Pach, Prof.	AGH University of Science and Technology
Wojciech Palacz, PhD	Jagiellonian University
Bartłomiej Pokrzywka, PhD, Associate Professor	Pedagogical University in Cracow
Marek Policht, MSc Eng.	Institute for Advanced Manufacturing Technologies
Tomasz Szmuc, Prof.	AGH University of Science and Technology
Jacek Barski, Deputy Director	Marshal Office of the Małopolska District
Grzegorz Żych, Director	Cracow City Hall

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.

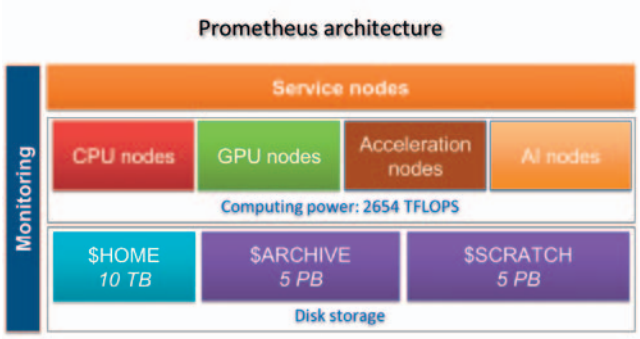
Prometheus in numbers	
Number of computing cores	53 748
RAM	283,5 TB
Number of GPGPUs	144
Computing power	2.65 PFlops
TOP500 – the list of the world's fastest computers (June 2020 edition)	288th position

Prometheus once again has been listed on the TOP500 list of the world's fastest computers (June 2020 edition) and took the **288th position**, again as the only supercomputer from Poland.

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.**

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

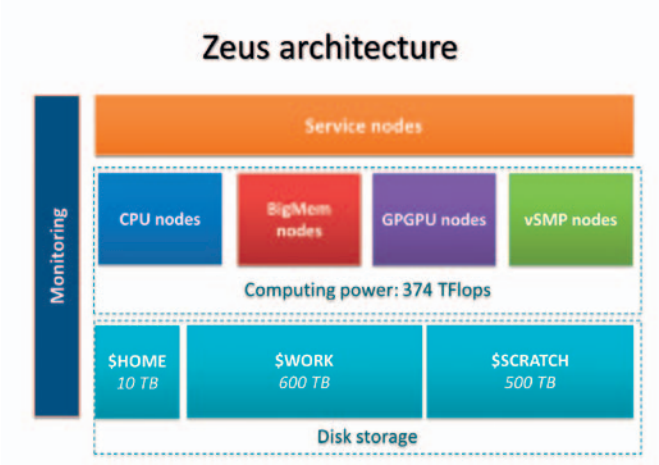
ZEUS – almost 60 000 CPU-years

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

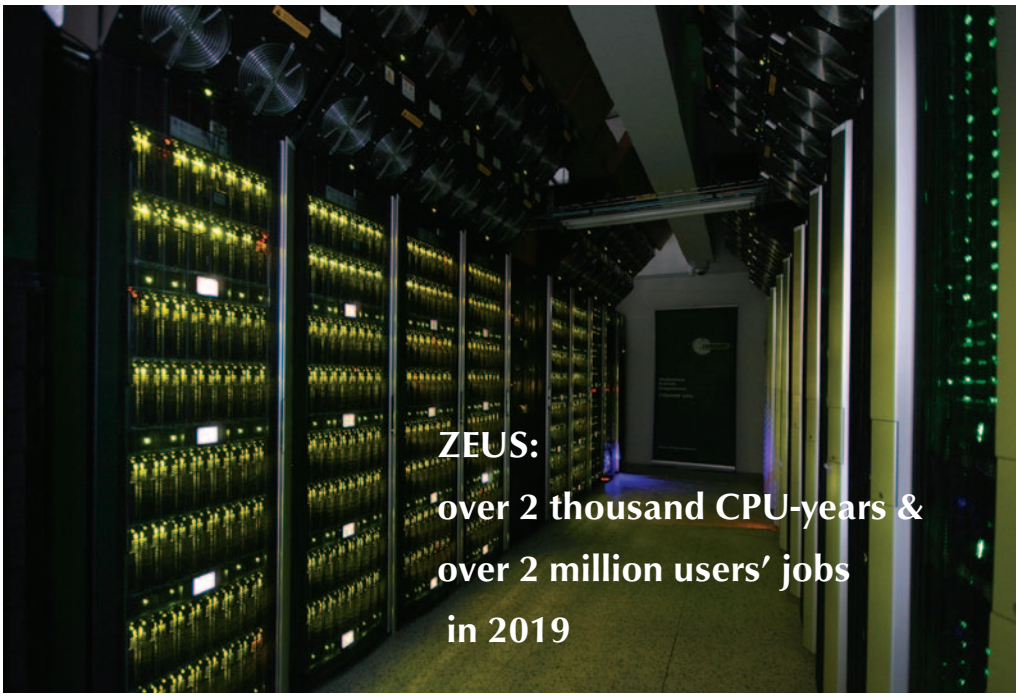
Zeus is a heterogeneous computing cluster. It constitutes of four classes of nodes, varying in terms of architecture of computing resources, specifically tailored to the requirements of the scientific communities. The Zeus architecture is a composite of four partitions:



- **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 2019 alone Zeus performed almost **2 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

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:

- modeling the activity of selected anticancer drugs through carbon nanoparticles,
- simulation of bio-based materials for drug encapsulation,
- analysis of DNA sequences preferentially recognized by human transcription factors YY1 and YY2,
- modeling of selected proteins, their complexes and complexes they form with tRNA and small molecule ligands,
- development of a numerical model of heat distribution in the Earth's atmosphere,
- research on electron structure and ion transport in oxide materials,
- superconductivity in selected physical systems,
- research on electron and transport properties of new magnetic materials.

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.

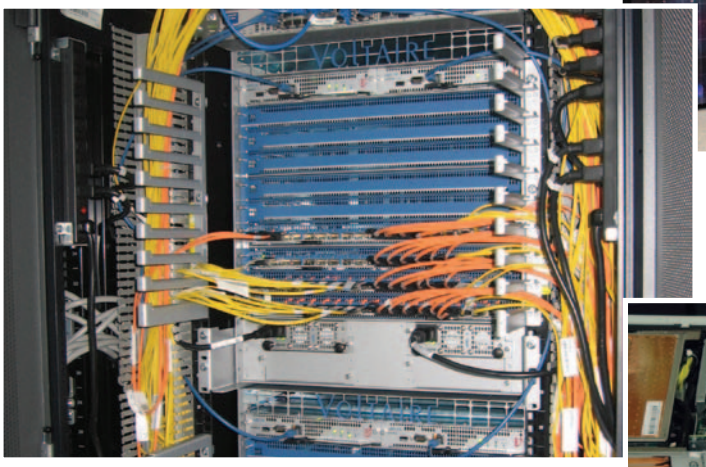
Obviously, even the highest positions in the TOP500 list, or the latest technologies used to build high-performance computers do not fully reflect the importance of this kind of computing resources for

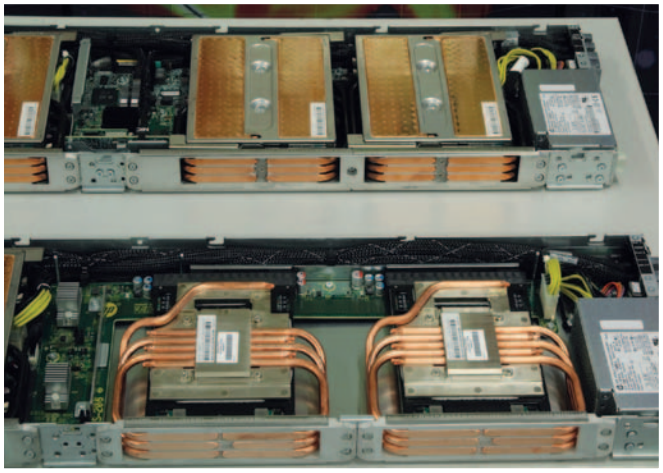
the Polish scientific community. The usefulness of supercomputers provided by ACC Cyfronet AGH as a tool for conducting research is best evidenced by statistical data on their use.

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
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





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.

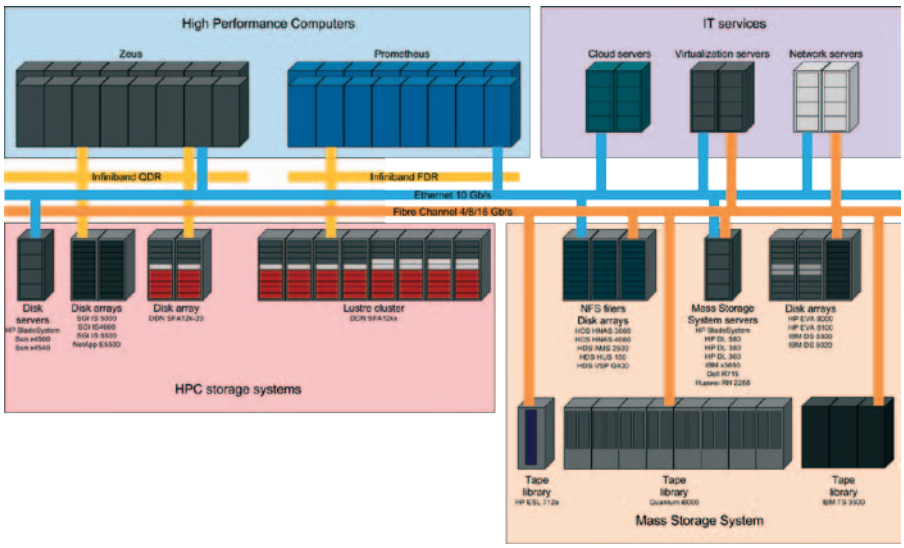
The most fundamental is the one used for keeping users' home directories. In this case all the crucial elements provide a very high level of availability and data security, which are supported by mechanisms such as snapshots and backups to external tape libraries. Zeus and Prometheus (the two main supercomputers 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 **49 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 **24 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.

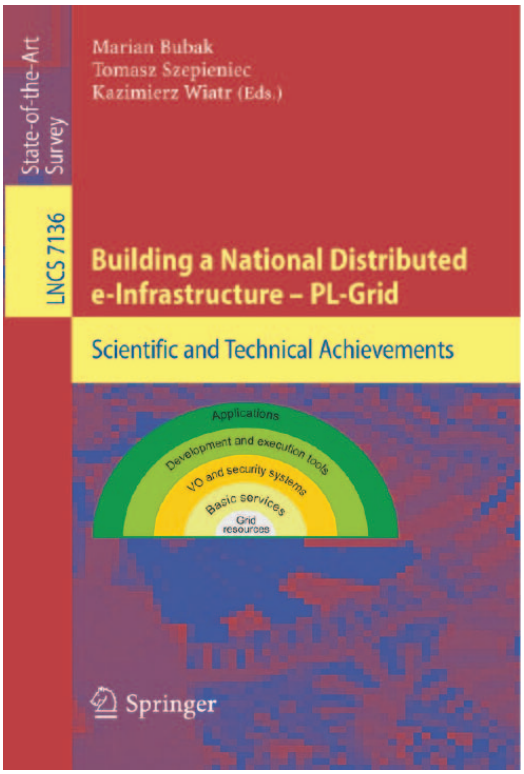
In the PLGrid NG project, the domain-specific grids were developed for several other groups of scientists, representing fourteen research fields (**in total, in the two projects, IT support tools were built for 27 scientific disciplines**).

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!**

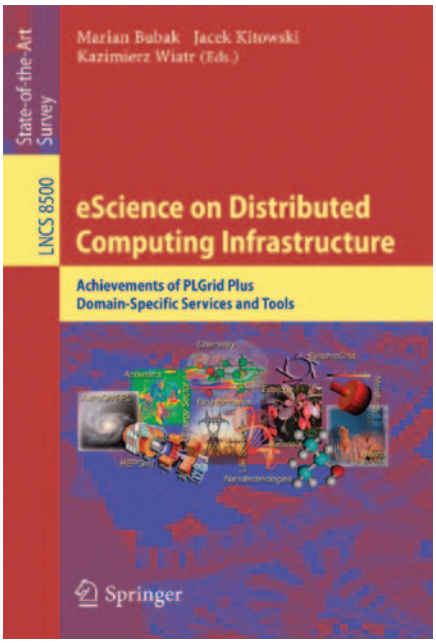
At present, more than 5 PFlops of computing power and more than 60 PB of disk storage are available within the infrastructure. In addition, many tools supporting organization of computational experiments, designing and running applications, computationally supporting research and results' visualization 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.

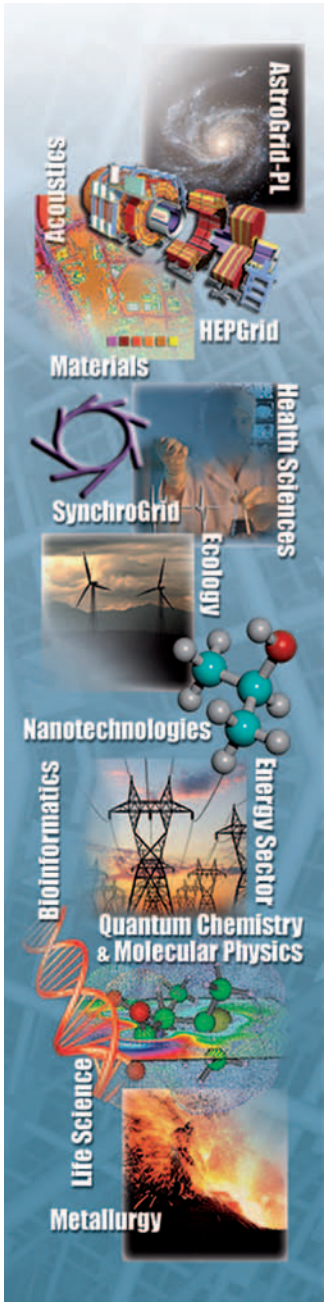
The second part (chapters 9 to 13) presents new environments and IT services that can be used by all of the previously mentioned groups of scientists. The third part (chapters 14 to 36) describes how specific environments, tools and services, prepared within the PLGrid Plus project, are used in advanced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

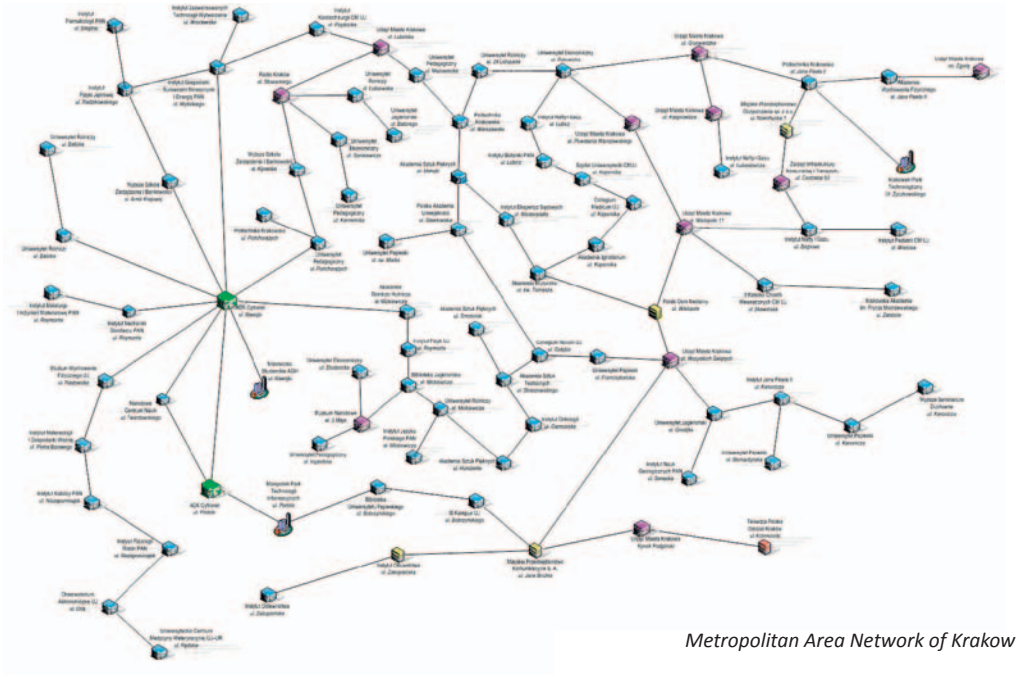
Preparation and implementation of a set of domain-specific services fit very well with the need of development of an advanced IT infrastructure designed for the implementation of modern scientific research. The well-tailored PLGrid e-infrastructure does not only fulfil researchers' needs for suitable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



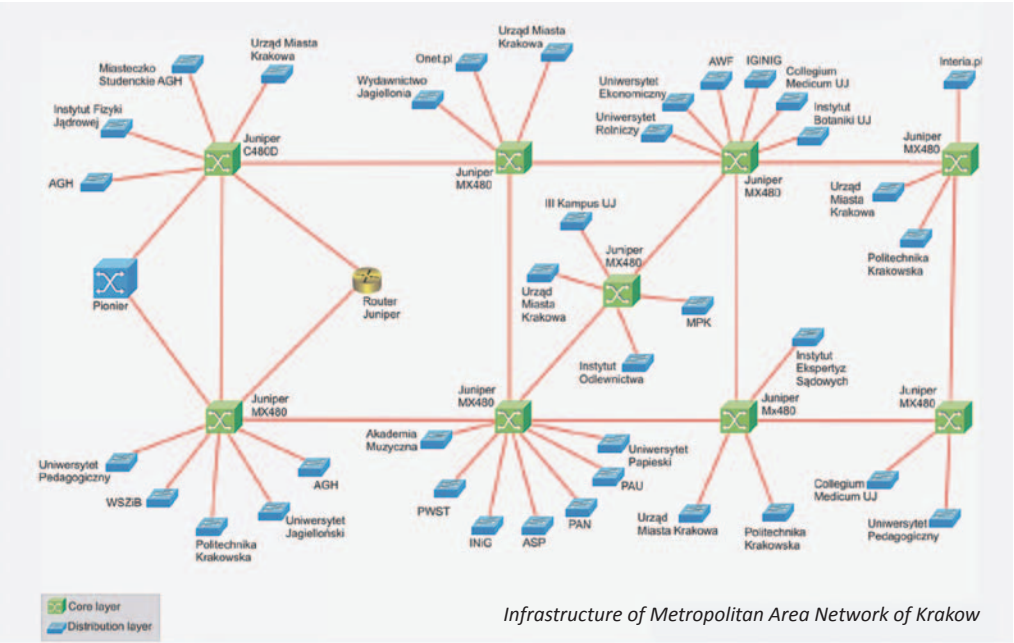
Metropolitan Area Network of Krakow

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 almost 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 4 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:

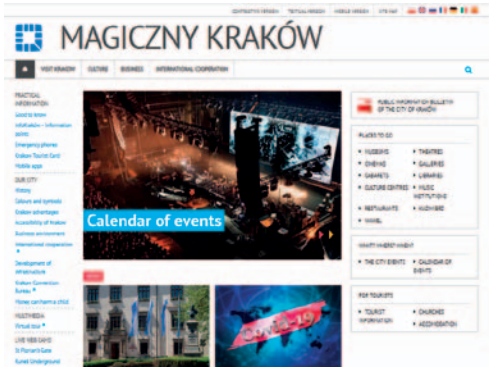
- **e-mail** accessed via SMTP protocol or web interface *http://poczta.cyfronet.pl*,
- **www**: CYFRONET operates a set of web sites, which in addition to news from the world of science, present information on the culture, sights and many other fields,
- **news**: discussion groups covering all areas 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,
- **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 2019	
Number of e-mails	~ 16 500 000
Number of e-mail server sesions	~ 41 000 000

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.

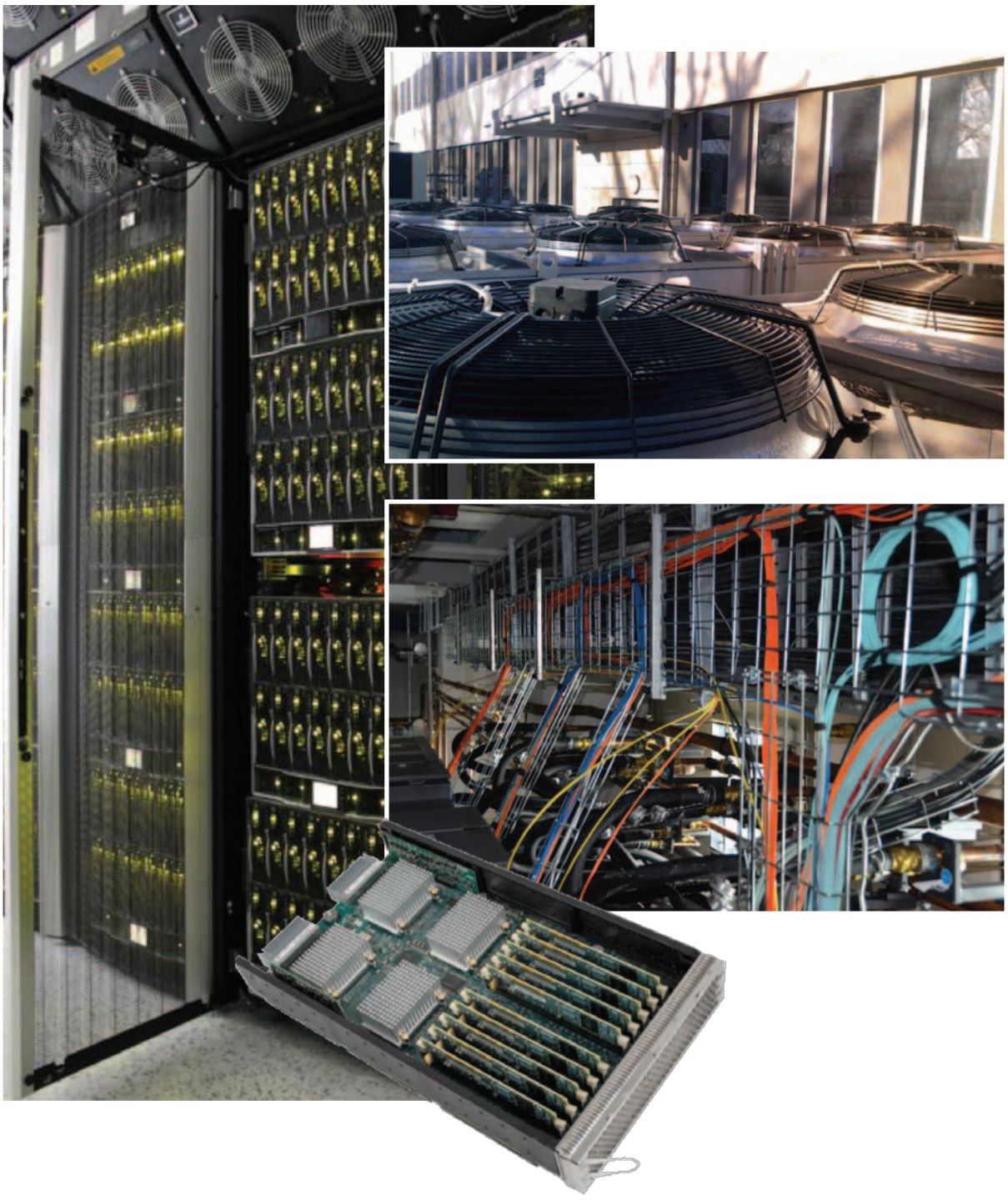
Cooperation with Kraków authorities is of particular importance for the Centre. The agreement between the Municipality of Kraków and CYFRONET, regarding the promotion of the City has resulted in the creation of an up-to-date portal. Aside scientific information the portal introduces its readers to the 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 and Zeus supercomputers provide:
3+ PFlops
79 000+ cores
300+ GPGPUs
340+ TB RAM

Storage

21 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

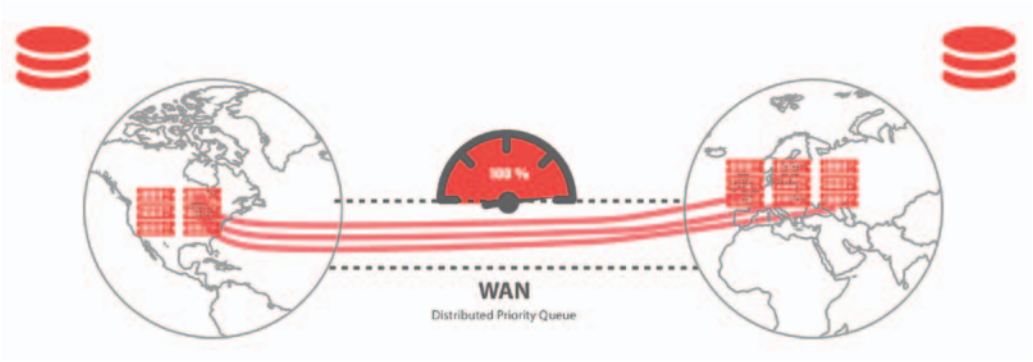
- Simple deployment based on Docker containers using a friendly command line client.
- Easy storage support for user requests based on secure tokens.
- Complex monitoring information available on all aspects of the system, accessible through REST API or directly visualized in the administration panel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



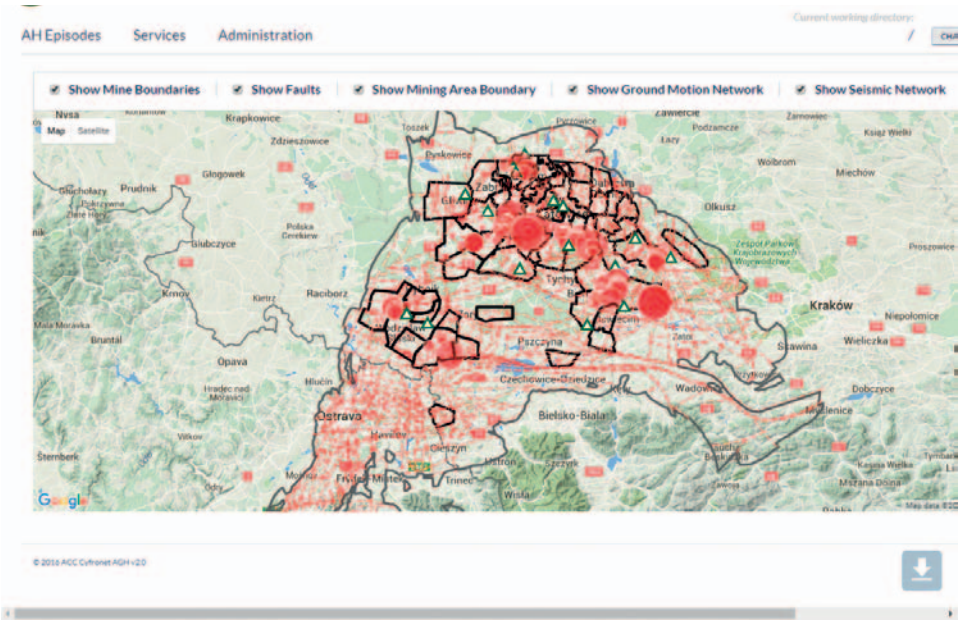
More information: *https://onedata.org*

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.

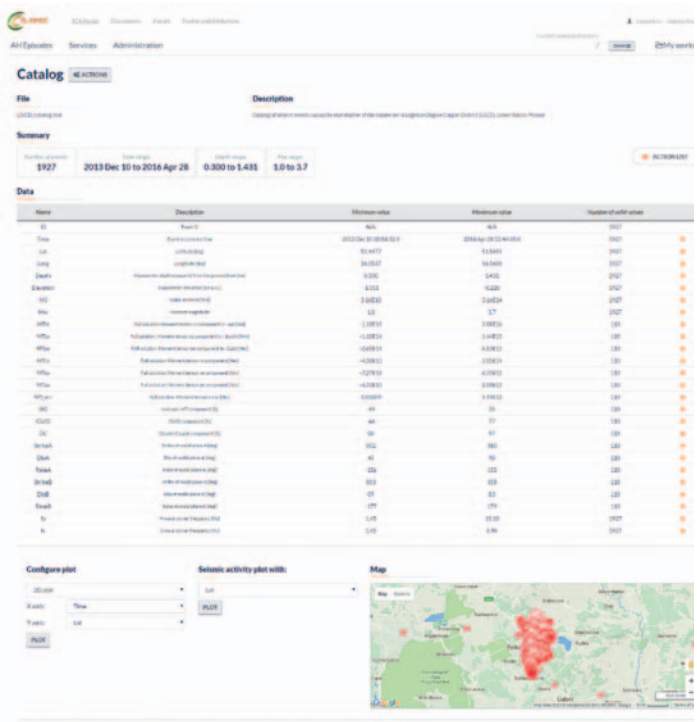
An InSilicoLab-based portal is designed as a workspace that gathers all that a researcher needs for his/her *in silico* experiments. This means:

- 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

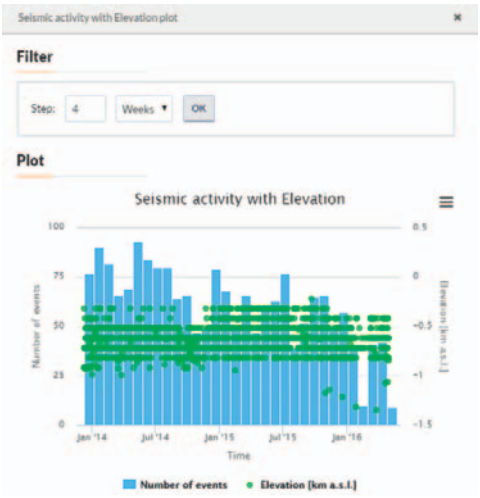
InSilicoLab is not meant to be a "*Yet another engine for...*", therefore, its developers has put maximum stress on the utility of the tool. This means user-friendliness, but, even more importantly, serving real scientific problems. This requires focusing on solving specific problems, rather than building 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.



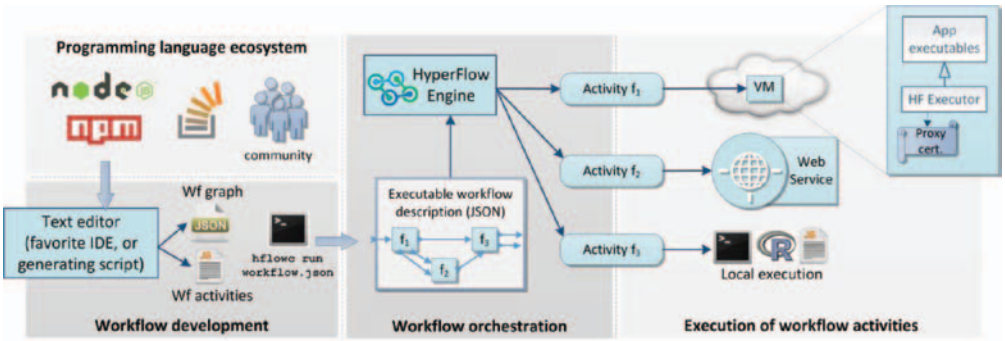
Domain applications

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.

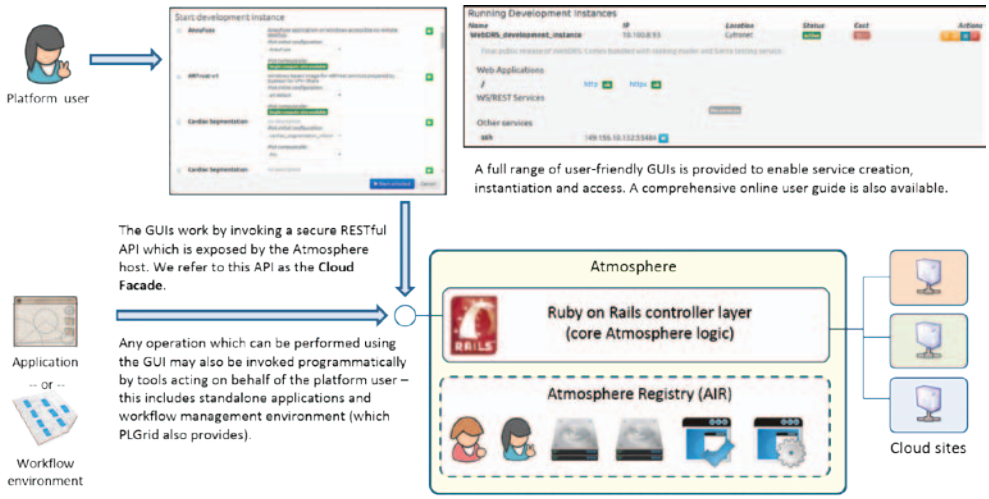
Contact

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.

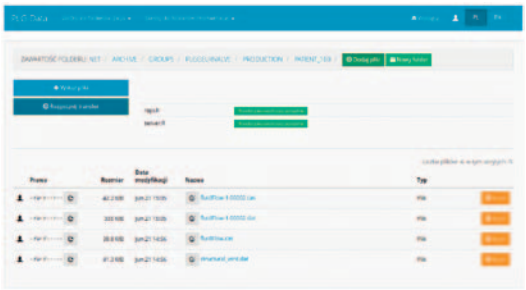
https://cloud.plgrid.pl

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 functionality 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.

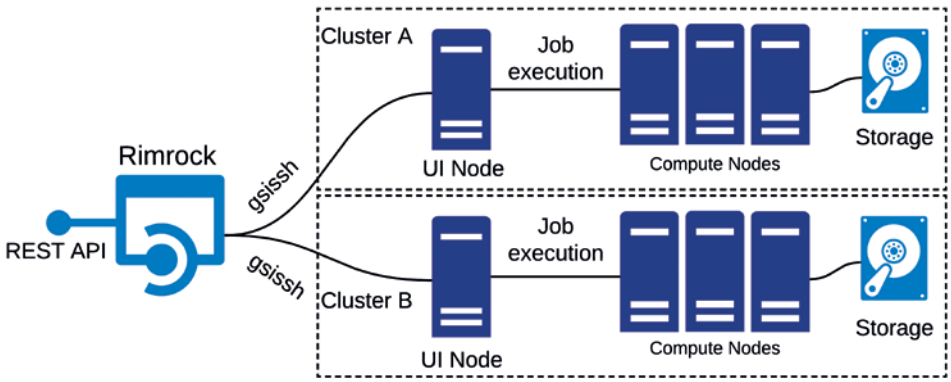
Logging in to the tool is done using either the PLGrid user-password pair, or a p12 certificate installed in one's browser. The service is available either in Polish or English. An advanced programmer's interface (API) helps developers to integrate their platforms, tools and services with the file storage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

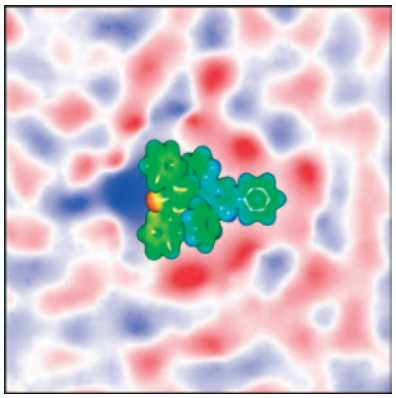
Data security

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:

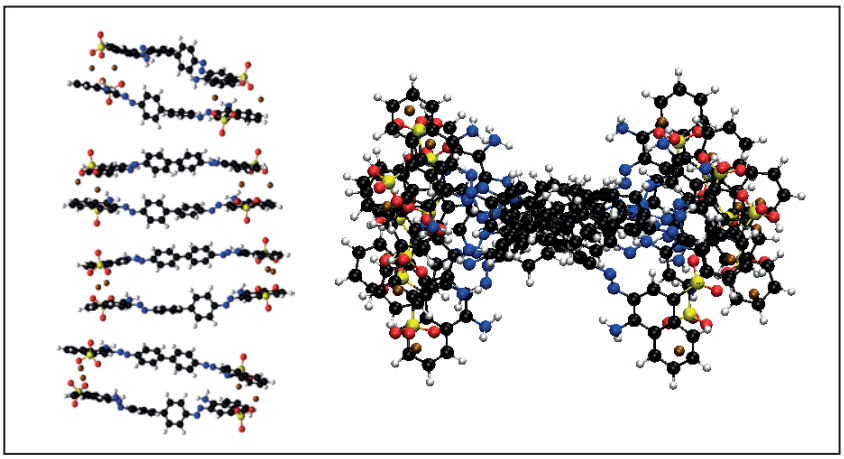
- Versatile and widespread used quantum chemistry codes such as **Gaussian**, **GAMESS UK**, **NWChem**, **Schrödinger**, **Q-Chem** and **TURBOMOLE**, which are capable of calculating electronic structure and various properties of diverse molecular systems using both *ab initio*, density functional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

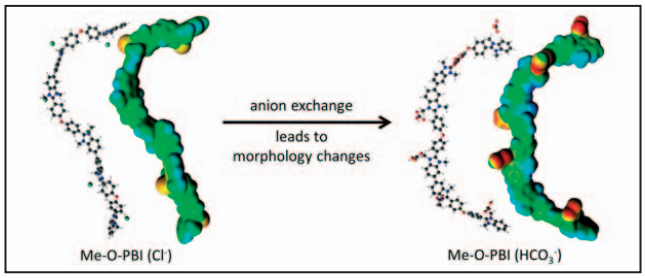
- Several packages, which could be used for **solid-state systems**. Among them **BAND**, **Quantum ESPRESSO** and **SIESTA** are worth mentioning.

- **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.

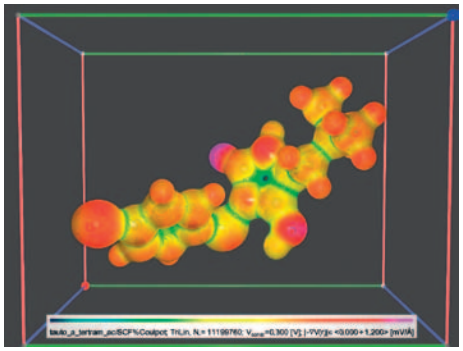
Scikit-learn is a software machine learning library for the Python programming language. It features various classification, regression and clustering algorithms including support vector machines, random forests, gradient boosting, k-means and DBSCAN. It has been designed to interoperate with the Python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

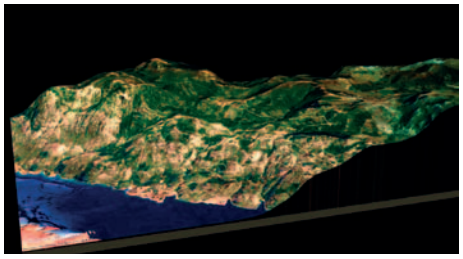
Data visualization enables analysis and understanding of the results of even very complex numerical calculations, especially multidimensional or time-dependent. Most applications for numerical calculations have a module that generates their visualization. Python has a matplotlib or VTK+ module, while 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:

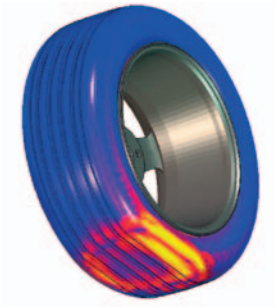


Janusz Opiła: Terrain visualization based on altitude data collected on an irregular grid and textures obtained from the Google Earth Pro application vicinity of Karlobag, Croatia). Own study: DOI: 10.23919/MIPRO.2018.8400037

- 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.

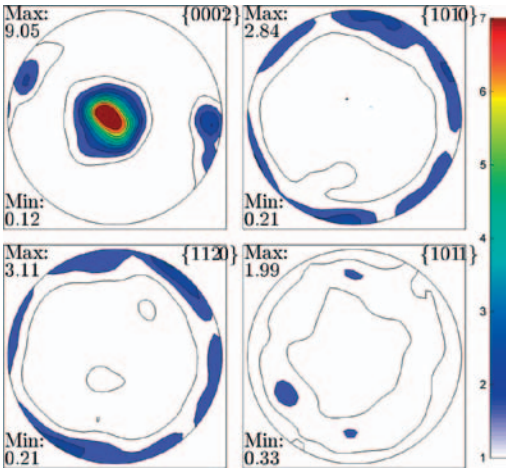
FLUENT software offers the broad physical modeling capabilities needed to model flow, turbulence, heat transfer and reactions for industrial applications ranging from air or liquid flow to semiconductor manufacturing. FLUENT can be used in numerous science domains, including chemistry, metallurgy, biomedicine,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

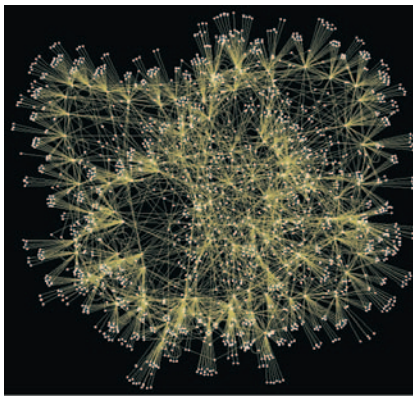


Bartosz Sułkowski: Results of texture simulations by visco-plastic self-consistent model of Zn after hydrostatic extrusion at 250 °C

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)

Laboratory of 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

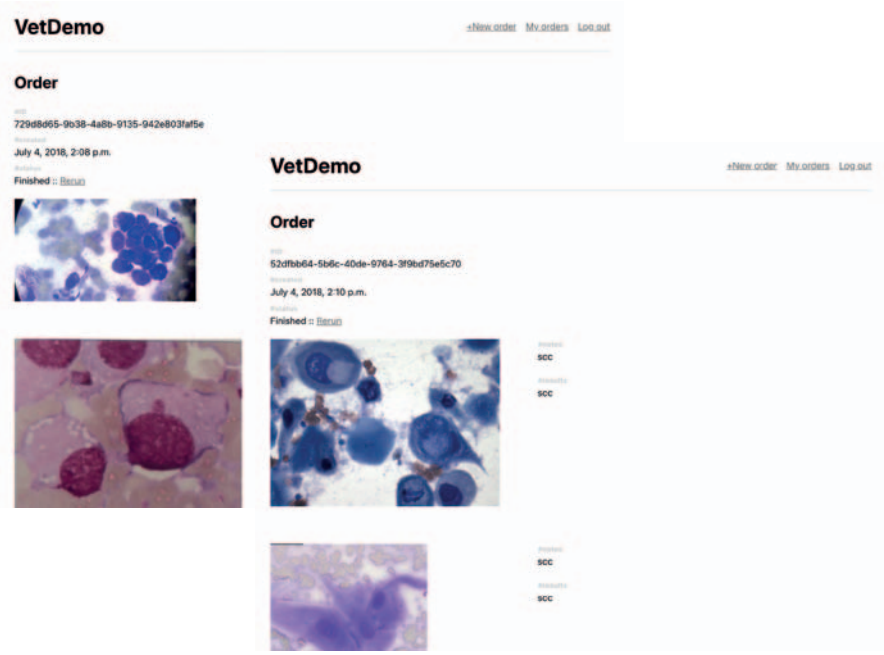
The Laboratory is proud of many years of experience in natural language processing. As an example, we have created and developed the tool that allows the users to search, compare and classify text documents. The result of this work is the web service called Scholar that is available at ACC Cyfronet AGH 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

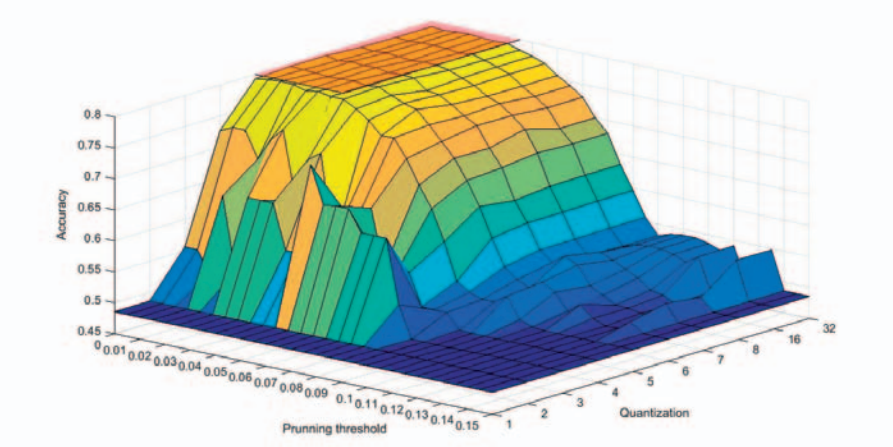
of the animal tissue. The proposed system achieved an efficiency of approximately 96% for the three selected types of cancer. We used deep network models based on Resnet-50 and Resnet-152 nets. Also, to allow for the selection of the perception area in classification operations, we have developed a special 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.



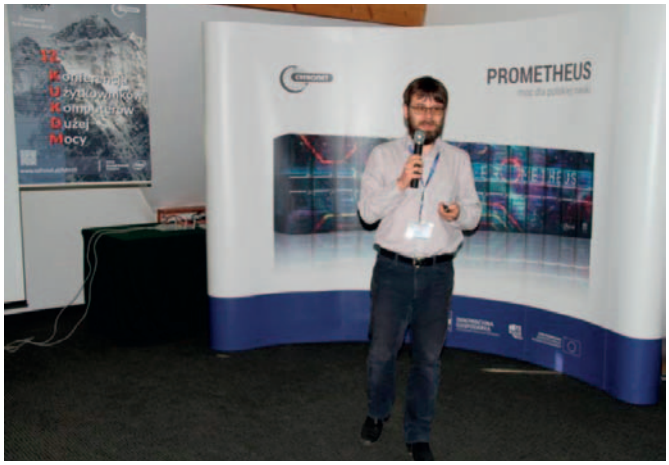
Time series analysis

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.

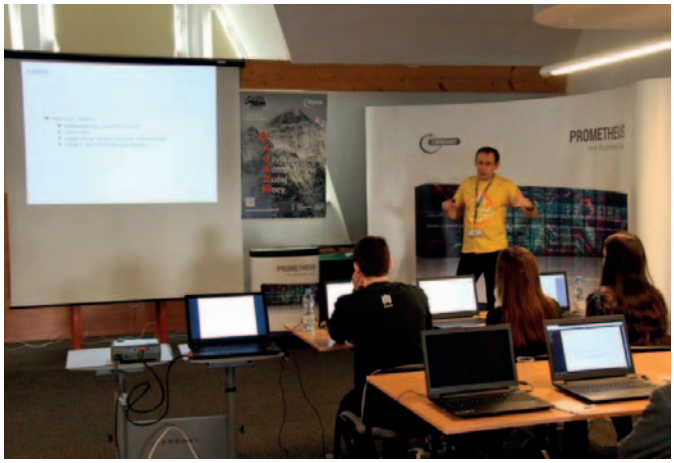
The conference was launched in 2008 and included several presentations by Cyfronet employees – describing the resources available in the Centre, as well as numerous lectures of researchers – presenting the scientific results achieved using Cyfronet hardware and software. In addition, two invited speeches 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.

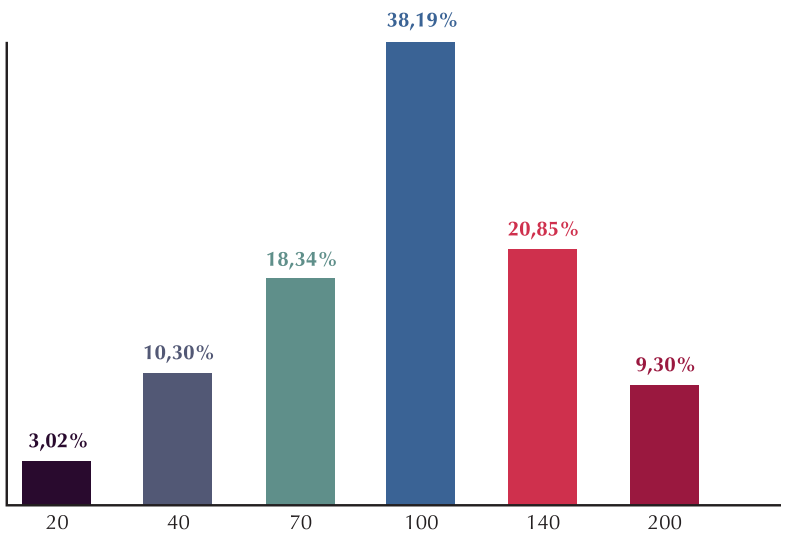
The crucial parts of the conference are meetings with suppliers of Cyfronet hardware and software, as well as the panel discussion on efficient use of these resources. The latter is always attended by users – researchers, who use the chance to get familiar with news regarding the computing infrastructure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



The conference is accompanied by poster and training sessions – in 2019 the training was focused on the programming of scientific calculations in the Python language.

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*).

http://www.cyfronet.pl/kdm19/



Ministry of Science and Higher Education marks of articles published in 2019 by Cyfronet Users in scientific journals

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. This international research foundation is one of three Polish beneficiaries of the prestigious *Teaming for Excellence Phase 2* call, as well as the only participant representing the Małopolska region.

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** 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.

An important aspect of the activities of **Sano** is its collaboration with the University Hospital in Kraków and its personnel.

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

European Open Science Cloud (EOSC) is an initiative aiming to create 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

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.

The second of the currently implemented projects, EOSC Enhance (https://www.eosc-portal.eu/enhance), aims to improve the Portal in terms of convenience and speed of use. The works carried out by our specialists concern, among others, advanced analysis of user behavior in order to create and implement the best user experience practices. Thanks to research, interviews and gathering feedback from real users, the Portal is developing, becoming more and more user-friendly. At the same time, new functionalities are being implemented to make it easier for researchers to use European research services and open science datasets.

Both EOSC-Hub, EOSC Synergy and EOSC Enhance are funded by the European Union's Horizon 2020 program.

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 EOSC data sharing policy is based on the FAIR principles: findability, accessibility, interoperability, and reusability. Thanks to this approach, Open Science from a theoretical paradigm becomes the reality.

EuroHPC – towards exascale computing

The European High-Performance Computing Joint Undertaking (EuroHPC JU) was established to radically develop 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 32 countries as well as private members with the aim to buy and deploy two exascale supercomputers that will be on the TOP5 list of the world's fastest computers.

The LUMI consortium

To achieve the EuroHPC goal, firstly pre-exascale machines will be operated. Poland, represented by ACC Cyfronet AGH and Ministry of Science and Higher Education, has joined the Finnish-led LUMI consortium, which works on the machine that will be ten times more powerful than the most powerful supercomputer currently in Europe (HPC5 from Italy). 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

To meet the growing need for computing resources of academia and industry, the LUMI supercomputer is designed to have computing power exceeding 200 PFlops (0,2 EFlops), which will be achieved by a combination of General Purpose Processors (GPP) and GPGPU partitions. Jointly with over 60 PB 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 e.g. astrophysics, weather prediction, seismic activity, personalised medicine.

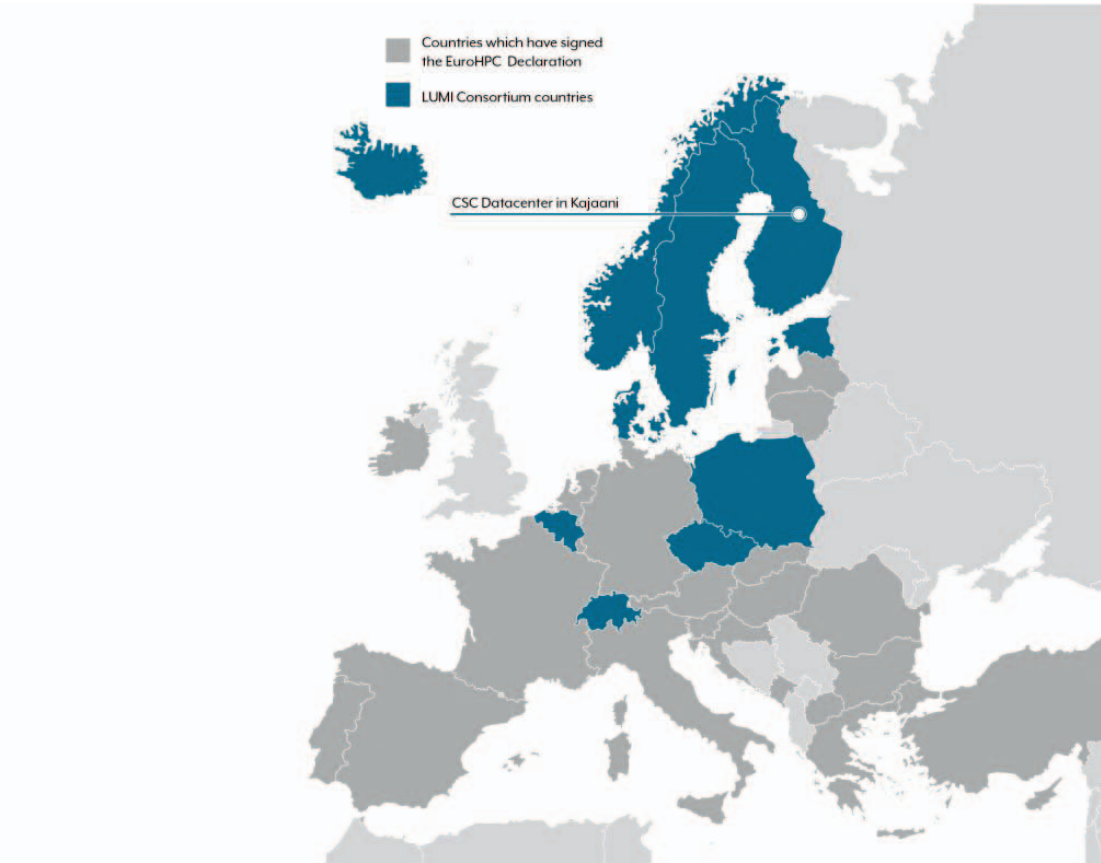
"We will be able to use the supercomputer to solve the problems that are unsolvable, taking into account the currently existing machines – says Marek Magryś, Cyfronet's Deputy Director for HPC. – *The flagship example are brain simulations."*

At the time of installation, the LUMI supercomputer will be one of the world's fastest computer systems. It will be operated at the CNC data centre in Kajaani (Finland), but its computing power, storage and services will be available for Polish scientists. One of Cyfronet's goals is to provide them with an efficient and user-friendly working environment with this supercomputer, based on good practices developed and national infrastructure adapted to new conditions.

Carefully selected site

The choice of Kajaani data centre was preceded by a thorough analysis of possible sites, as they should fulfil the strict EuroHPC criteria. CNC data centre is situated in Northern Finland and due to the climate it benefits from natural cooling provided all-year-round. Kajaani is also equipped with three local hydro power plants (30 MW) and a biomass power plant (88 MW) which guarantees green energy delivery. And the waste heat from the supercomputer will be utilized to warm data centre and surrounding buildings, what will reduce the negative impact on the climate.

The LUMI supercomputer is expected to be launched in the first half of 2021.



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.*

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 aims to build an improved, more integrated version of the EOSC Portal, which will enable improvement and extension of solutions that will 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 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 multiscale 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 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 EPOS Implementation Phase project (EPOS IP) built on the achievements of the successful EPOS preparatory phase project (EPOS PP). The EPOS project was integrating the diverse, but advanced European Research Infrastructures for solid Earth Science, and built on new e-science opportunities to monitor and understand the dynamic and complex solid-Earth System.

The aim of the project was to develop a comprehensive, clinically-compliant decision-support system to meet the challenge of treatment optimisation in case of the Valvular Heart Disease, by quantifying individualised disease severity and patient impairment, predicting disease progression, ranking the effectiveness of alternative candidate procedures, and optimising the patient-specific intervention plan.

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 aim of the project was the development of a model-based predictor system, supporting the flexible design of strip rolling, joining functionality of numerical simulations, material modelling, sensitivity analysis and optimization.

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 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 is digitization, preparation of descriptions of digitized data and the final publication of digitalisers on the PAUart platform. 13 140 new records will be published on this platform, including graphics from the modern period, valuable in artistic and scientific terms.

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.

Within the project a comprehensive diagnostic test was developed, based on analysis of the material obtained from the tumor via surgery or biopsy.

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

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

The project aimed at integration of new groups of researchers from the next 14 disciplines with the PLGrid computing infrastructure. Thanks to deployment of new domain grids, research teams from these areas are able to faster obtain results of their calculations and to better integrate with the national computing infrastructure for science.

The project supported, by means of IT, Polish research teams in conducting research and also enabled extensive collaboration among these teams, as well as international cooperation in the area of e-Science.

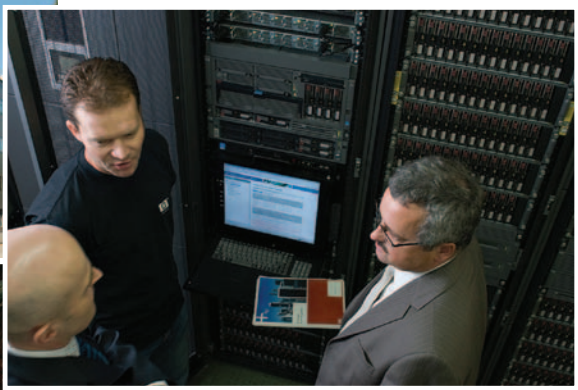
The preparatory phase of Polish input for Cherenkow Telescope Array Project. This project has been undertaken to help design high level tools easing CTA data analysis. The main area of work covered the development of InSilicoLab Science Gateway for CTA with focus on job results analysis and visualisation, in particular.

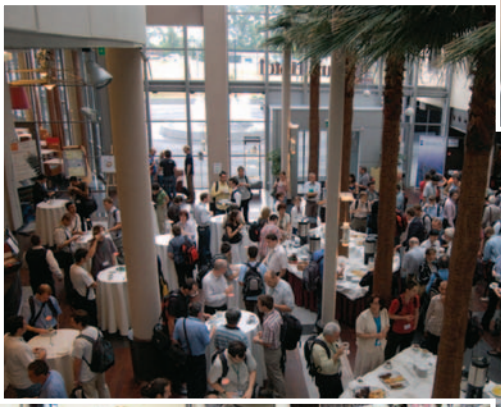
The aim of the project was integration of selected services available in the PIONIER network, and the development of the new services, e.g., with the increased reliability and security.

The 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.

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.

Within the project the Polish Grid Infrastructure (NGI) has been built to provide the Polish scientific community with an IT platform based on computer clusters, enabling research in various domains of e-Science.






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 2021 Contest edition!

http://www.cyfronet.pl/konkurs

The laureates of one of the previous Contest editions

Mateusz Sitko

The interview with the author of the PhD thesis: "Development of computationally efficient cellular automata model for recrystallization"

What made you start modeling metal structures? What aspects are most interesting to you?

The most interesting aspect of my work is the possibility of conducting research at the interface between two fields of science, i.e. the use of knowledge in the field of computer science and materials engineering in order to develop and implement systems supporting the development of new materials, as well as technologies for their shaping.

Already during my engineering and master's studies, I became interested in the development of applications for computer-aided design of advanced metal alloys shaping technology, using the concept of a digital twin for material. This work allowed me to get to know the issues related to the concept of digital representation of materials, which gives the possibility not only to reproduce what is happening in the material at an unprecedented level of accuracy, but also to predict the behavior of the material in increasingly complex states of stress. When working with materials such as modern metal alloys, the most interesting part is the practical side of problem solving and the applicability of the solutions obtained.

How would you introduce the use of cellular automata to simulate recrystallization to someone outside the field?

Imagine the situation of designing a new structural element, for example for a car, which must meet the exorbitant expectations of engineers in order to ensure an adequate level of user safety at an acceptable cost. Obtaining such a finished product is associated with a number of different plastic working and heat treatment processes. If we want to obtain a product with a precisely defined shape and the mentioned functional parameters, it is necessary to strictly control many process parameters such as time, temperature, or the method of cooling or heating. These parameters directly affect the changes taking place in the material at the microstructure level, which in turn directly affects the functional properties of finished products. Therefore, controlling all such micro-scale changes, e.g. remodeling of the microstructure as a result of static recrystallization, is particularly important from the point of view of developing a new production technology for products with the required properties. Cellular automata are one of the few methods with which it is possible to directly map, and thus control, what happens to the microstructure of the material during thermo-mechanical processing.

What was the greatest challenge in research involving metallurgical theory, numerical modeling, programming and experiment?

The greatest challenge that arose during the work was laboratory research, carried out during an internship at the Deakin University in Australia. For the purposes of my doctoral dissertation,

I designed and carried out a plastic deformation experiment along with various heat treatment cycles and subsequent metallographic analysis. As on a daily basis at work I mainly deal with the implementation of various types of algorithms, it was a big challenge for me, but at the same time it significantly broadened my understanding of the material behavior, I worked with.

Can you indicate milestones of work on a doctoral dissertation?

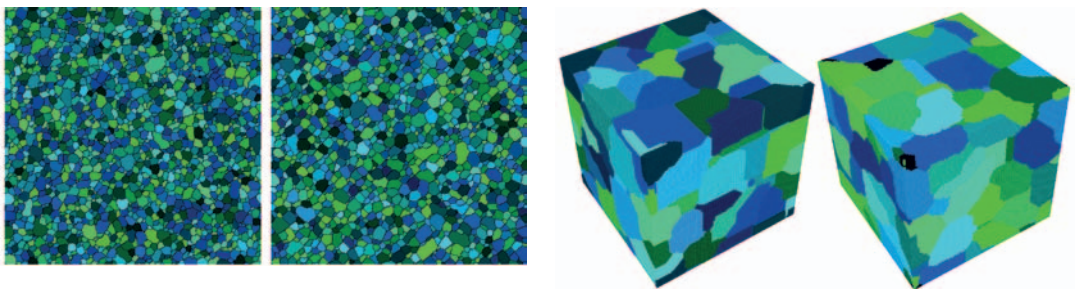
The first milestone of the doctoral dissertation was the development and implementation of a comprehensive model of cellular automata of microstructure development during static recrystallization, taking into account the most important mechanisms occurring during this phenomenon. The second key stage of the work was the aforementioned laboratory tests, which allowed for the identification of the parameters of the developed model and its subsequent verification. The last and most important aspect was the adaptation of the developed algorithms to their parallel execution. The concept of parallelization uses a problem-domain decomposition approach for application in multi-node systems using MPI technology. During this stage of research, algorithms responsible for sending messages were developed, and problems related to the overlapping of calculations and communication were solved. Various communication mechanisms between the computing domains were also proposed and implemented.

How did you use the resources provided by Cyfronet? What opportunities turned out to be the most valuable?

Cyfronet resources allowed for the execution of calculations in very large, discrete computing spaces. Processing such data on classic servers would not be possible both in terms of the availability of operational memory and computing time. I made calculations on the Prometheus supercomputer, among others for digital twins of material microstructures obtained in collaboration with the Los Alamos laboratory in the United States, using a circular particle accelerator.

What tips could you share with people just starting their doctoral studies?

The most important advice that I can give to younger colleagues is to find the right team with passionate people. A doctorate is work for oneself and if it is carried out in a friendly atmosphere, then solving more and more difficult problems will be really rewarding. Even if someone, after completing the doctorate, does not think about working at a university, then by carrying out research under international projects, one can gain not only very good experience, but also establish cooperation with the best foreign centers, and polish the language during foreign internships. The second piece of advice I can give is not to be afraid of risk and take up scientific challenges that will actually allow you to significantly push the boundaries of our current knowledge!



Example of 2D and 3D digital material representation

Wojciech Szczypka

The interview with the author of the PhD thesis:
"*In silico* studies of electronic structure and bonding properties of selected thermoelectric materials from ternary chalcogenides and silicon clathrates groups"

How did you get interested in computational chemistry?

During classes in general chemistry in the first year of college, I often solved tasks at the blackboard, using the skills acquired during many high school competitions, including the Chemistry Olympiad. My solving methods, due to the extensive mathematical apparatus, caught the attention of the teacher, who soon instructed me to contact another employee of the Faculty, specializing in computational chemistry. In this way, I came under the care of my long-term research tutor and supervisor with whom I have worked for almost eight years. A simultaneous passion for mathematics, chemistry and physics as well as a fascination with computer technologies made me love computational chemistry, which poses challenges in each of the above-mentioned fields.

What made you devote your doctoral dissertation to specific compounds from the chalcogenides and silicon clathrates groups?

Before I even started thinking about doctoral studies, I was already heavily involved in theoretical calculations for thermoelectric materials. Even before planning the application for the Diamond Grant, which I later received, I often spoke with experienced scientists in the field, both at the Faculty and during scientific conferences. In this way, I selected particularly interesting systems that were worth investigating with methods popular in computational chemistry (electron density topology analysis, bond valence methods), and rarely found in research work on thermoelectric materials, in which the analysis typical for solid state physics definitely dominated. In the case of chalcogenides, materials mostly well-known both from the experimental and theoretical point of view, I chose $AgSbTe_2$ and the ordering in the cationic subnetwork as well as the influence of the presence of intrinsic defects. Although both of these factors significantly affect the properties of the material, little research work has been done in this area so far. Silicon clathrates are materials that differ significantly from chalcogenides in terms of chemistry. Their choice was related to the willingness to demonstrate in my work the universality of the methods used, allowing to understand many important features of both groups of materials. In the case of silicon clathrates, I took one type of structure out of many possible for these materials, for further cross-sectional analysis in order to reduce the complexity of the tested systems.

What were the biggest problems during the research and how did you overcome them?

When planning research for my doctoral dissertation, I initially collaborated with a group of experimenters, hoping to jointly publish scientific papers. Unfortunately, in the meantime, the other

side's plans had to be changed, and eventually I had to deal with it myself. For this reason, I needed to delve even more into the literature already published, in order to adapt the research plan to purely computational work, at the same time focused on specific and current problems concerning the studied groups of materials. Ultimately, this problem was the driving force for me to ensure that my research was world-class compared to other theoretical work. Together with the supervisor, I was able to publish my results in very good journals of the Philadelphia list.

How did the resources provided by Cyfronet support your work?

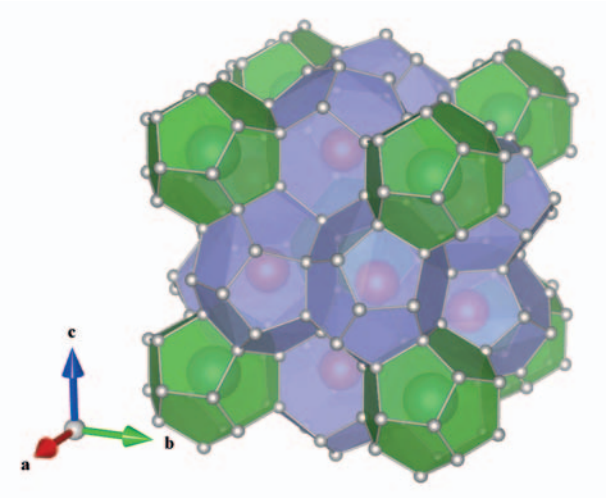
Actually, apart from small calculations in spreadsheets, almost all research, both from the time of work in the student research club and later during doctoral studies, were calculations made on the resources of Cyfronet. At first we used the Zeus supercomputer, and then we moved to Prometheus. I used software licensed by Cyfronet and compiled open source programs using the prepared environment; sometimes in periods of low traffic I could run many difficult computing tasks, often using several hundred cores simultaneously. In total, I used over 750,000 computing hours!

How do you assess the directions of research development on thermoelectric systems and what, in your opinion, is needed to create a large-scale installation?

Research on the possibilities of using thermoelectric materials in modules for the production of electricity has been going on for decades. Thermoelectric generators have been successfully used in space probes, among others due to the lack of moving parts, and in many other specialized applications, in which they are often irreplaceable. However, the main factor limiting the use of popular thermoelectric modules on a large scale is their high cost, resulting from the rare chemical elements used in their production, which are often also highly toxic. Hence, the directions of research that seek materials devoid of these negative features are developing very dynamically. The efficiency of thermoelectric modules achieved so far is still insufficient to make them profitable on a large scale in such areas as automotive or energy. However, the rapid development of research observed in the last decade gives great hope that it will soon be possible to disseminate devices that are both economically and ecologically justified.

What could you advise people who are just starting their doctoral studies?

In my opinion, the most important thing is to be well prepared – to constantly expand your knowledge and your workshop, regardless of whether we conduct our research in a laboratory test tube or on a silicon plaster. You should also be aware that both methods are absolutely irreplaceable nowadays and when they interpenetrate and complement each other, they really enable crossing new boundaries in the study of reality. Therefore, it is very important to have a fundamental literature review, both cross-sectional and detailed, as well as an overview of the available methods and tools – all this will help create a good research plan. I have the following rule: "if you have to cut a tree in 5 minutes, take the first 3 minutes to sharpen your ax".



Structure of type I silicon clathrate with marked T_{20} and T_{24} frames

Karolina Śliwa

The interview with the author of the PhD thesis: "Plant micellar extracts as active ingredients of preparations for the care of atopic skin"

What prompted you to dedicate your doctoral dissertation to researching the ingredients of atopic skin care products?

From the beginning, I wanted my doctoral dissertation to be of practical use. When I started a literature review on atopy in 2010, I realized that there was a rapid increase in the incidence of atopic dermatitis (AD) – this problem affects over 20% of the population. I was most touched by the fact that atopic dermatitis is one of the most common skin diseases among infants and children, and in case of most patients the disease remains for life. Moments later my nephew was born, who was diagnosed with atopic dermatitis. Then I realized how difficult it was for him, his mother and the whole family. In the treatment of atopy, dermatologists prescribe ointments with glucocorticosteroids, but unfortunately their long-term use is associated with the occurrence of many side effects. That is why I decided to start looking for alternative care based on natural, non-allergenic ingredients and create an innovative series of effective products. There are currently no products of this type on the cosmetic and pharmaceutical market.

How would you introduce the Micellar Assisted Extraction process to a person unrelated to the field?

The process of ultrasonic assisted micelle mediated extraction (UAMME) is the extraction of individual chemical compounds or their groups from plants with the use of aqueous surfactant solutions. In order to increase the efficiency, the process is carried out in an ultrasonic bath. Surfactants are chemical compounds whose structure can be divided into water-soluble and water-insoluble parts, and thanks to that in water their molecules accumulate into larger clusters, i.e. aggregates or micelles. During the extraction of UAMME, active substances which under normal conditions are poorly soluble in water, e.g. flavonoids with antioxidant activity, are solubilized, i.e. dissolved, inside surfactant aggregates. Since in my research I used cosmetic surfactants, the extracts obtained by me can be directly used as raw materials for the production of cosmetics.

What are the main advantages of this type of extraction compared to the methods commonly used in the cosmetics and pharmaceutical industries?

Micellar extraction takes place at a temperature lower than extraction with boiling organic solvents, ensuring the stability of substances sensitive to elevated temperature. Economic considerations are also important – surfactant concentrations are low, and their presence in the final product may increase its value. The lack of contact with organic solvents also allows for wider use of the separated substances in direct contact with the human body. In the extraction of active substances by the UAMME method, non-ionic surfactants are most often used, because they exhibit good solubilizing properties and do not cause toxicological and dermatological problems. Moreover, the structure of micelles protects the extracted compounds against the harmful effects of oxidants and prolongs their durability. There is a potential possibility of composing such prepared extracts into various cosmetic forms, in which the surfactants introduced with the extracts can also fulfill their traditional roles: emulsifiers and wetting agents. For the cosmetics industry, the method is therefore waste-free, because the surfactants used are also a co-emulsifier in the emulsions obtained. The use of various micellar systems additionally allows increasing the selectivity in obtaining selected groups of chemical compounds, which may help in the elimination of potentially allergenic substances. It is a completely safe method, non-toxic for humans, environmentally friendly, and therefore belongs to the group of extraction methods based on the principles

of "Green Chemistry". The fact that speaks for the more and more frequent use of this method is primarily the high efficiency of isolated natural substances.

How did Cyfronet's resources contribute to the theoretical part of the research?

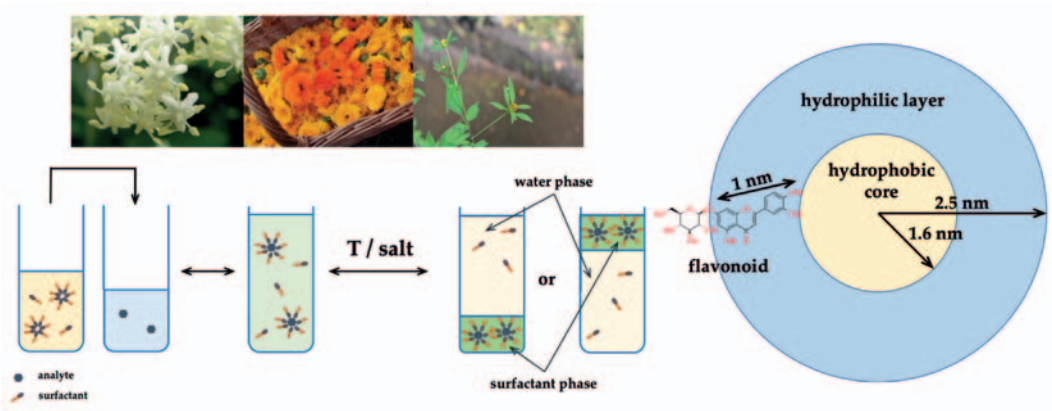
In order to analyze the phenomena occurring in the process of surfactant aggregation and the solubilization of flavonoids in a micellar solution, I have carried out theoretical research using the molecular dynamics method. A limitation of the applied methodology was the necessity to use surfactant models (the compounds used in the experimental part are most often mixtures) and approximations (simulation box size, short simulation time). Theoretical research at work turned out to be the key to explaining the mechanism of the process.

What further steps can be taken on the basis of your results?

Currently, four patents have been issued with recipes for preparations containing micellar extract from the three-part beggar-ticks. A pharmaceutical company is initially interested in one of the patents, and I hope this company will buy that patent from the Cracow University of Technology. I think implementation can be expected. I am also thinking about implementing my ideas, but I am absorbed by further scientific research. I am currently participating in a project on preparations for psoriasis. Besides, I am working on my own series of oncocosmetics.

What advice could you give to young scientists starting their research careers, based on your own experience?

Personally, I was not a doctoral student. I started my work at the university as a full-time technical employee, and at the same time I was given the opportunity to conduct research. Thus, since 2010, by conducting classes, scientific and research projects, going to scientific and industrial internships and scientific conferences, I had enough time to explore the topics of atopy, skin care, formulation, micellar-assisted extraction, as well as learning about new herbs, their composition and operation. As a young scientist, I believe that doctoral students at the beginning of their career, when they have an idea for an invention, should start the patenting procedure right away, because it is long-lasting. And unfortunately, publishing an innovation in the form of a doctoral thesis or publication will prevent them from obtaining a patent.



Extracting active ingredients from plants by micellar-assisted extraction

Dr Ewelina Wlaźlak

The interview with the author of the PhD thesis:
"Study of the influence of intermolecular interactions on the properties of selected iodides and triiodides""

Where does your interest in broadly understood electrochemistry, new materials, their properties and potential applications come from?

This interest is mainly due to the need for new materials. There are many ways to predict the properties of new compounds, but we usually only get clear information after receiving a new material. To understand what is actually going on in it, you need to study the electrochemical processes that take place there. You can then explain the properties of the new compound and verify whether we have obtained the material with the desired characteristics that is suitable for something, or another curiosity.

What were the biggest problems you encountered during the research and how did you manage to solve them?

During the research on the new photo catalyst I obtained a material that unfortunately could not withstand the conditions in which it should work (aqueous solutions). By changing the parts of the structure of this compound appropriately, it was possible to obtain a material that does not undergo hydrolysis ... but is subject to photo corrosion. The obtained material turned out to have memristive properties, which initiated an additional direction of research.

How did you use the resources provided by Cyfronet?

I used the resources of Cyfronet mainly to calculate band structures and electron densities of new compounds. The obtained results were extremely helpful in the analysis of interactions present in the structure of the new material, which have a dominant influence on the properties of the new semiconductor.

Based on the research results obtained by you and your team, what further steps can be taken?

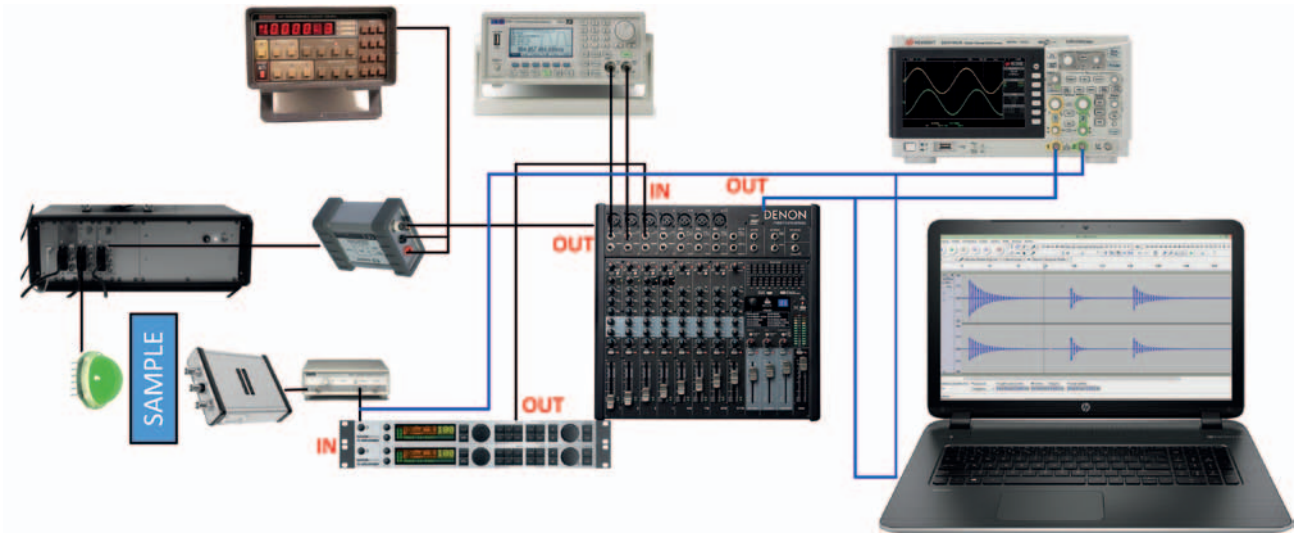
Most of my scientific work was the basic research and I do not plan any practical application of the compounds that I described in my dissertation. It is important that the results I obtained allow me to determine the key conditions that must be met by the material used to build a memristor, the operation of which would be based on the modulation of the Schottky's barrier. By following these guidelines, the material can be found among existing compounds that would be easier to implement and cheaper when deployed on a larger scale. However, this requires further research conducted a little bit differently than before.

What are your forecasts for the use of memristors for information storage and processing in the near future? Are we dealing with another technological revolution?

Companies such as IBM, HP or Samsung are working on the use of memristors. At the moment, it is difficult to say whether the materials used in memristors (mainly metal oxides) will withstand the stringent requirements and will be introduced on a larger scale. We are currently witnesses of a technological revolution in which memristors take part, but they are not its initiator. Every two years, the amount of information stored doubles, but only a fraction of it is analyzed. This phenomenon is often referred to as an „information black hole". Unconventional information processing seems to be the best solution to this problem at the moment. And here memristors used in hardware neural networks or reservoir systems have their place, among others thanks to the multi-state of this device. In addition to states 0 (say: off or high resistance memristor state) and 1 (on, low impedance state), it is also possible to use the entire range of states between 0 and 1, which opens up a full range of new possibilities.

What advice would you give to young researchers who are just starting their doctoral studies? What should these people pay attention to and what should they avoid?

First of all, a PhD student should choose a topic, he or she is interested in. Working on a project can sometimes take up much more time than expected, and nothing is more discouraging than working hard on non-addictive and unimportant issues. Second, it is sometimes worth taking on something that you may not be faintly aware of at first. The ability to deal with such situations is extremely valuable, and if someone has not yet acquired it, this is the last call. Thirdly, you should decide quite early what you are going to do after your doctorate. Contrary to what you may sometimes think – yes, there is life after a PhD, and you have to organize it somehow.



Reservoir circuit used in signal analysis using a memristor

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's users are the best proof of its usefulness as a tool for conducting scientific and research work.


photo: Michał Niewdana, NSC

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.

Award of the National Science Centre for Adam Rycerz, PhD DSc from the Institute of Physics of the Jagiellonian University*


photo: Michał Niewdana, NSC

On October 10, 2017, Adam Rycerz won the Prize in the field of technical sciences. The prize was awarded for theoretical analysis of quantum transport of a charge in graphene nanosystems. The jury considered it particularly important that the laureate had proposed a mechanism for controlling current polarization in the space of the valley degrees of freedom through electrostatic fields.

– *In the theory of condensed matter, I am particularly fascinated by the possibility of predicting – with relatively little effort and resources – completely new, previously unknown physical phenomena, the practical use of which is actually within reach* – says Adam Rycerz. – *Although the cases where the implementation of such a plan was successful, can be counted on the fingers of one hand*

(usually the experimental discovery preceded the laborious process of building the theory), abandoning such attempts will certainly mean the end of civilization.

Adam Rycerz supports his research with calculations on Cyfronet supercomputers within the CORRCO2 computational grant: "Macroscopic quantum effects and critical phenomena in Dirac materials".

Award of the National Science Centre for Joanna Sułkowska, PhD DSc from the Centre of New Technologies and the Faculty of Chemistry of the University of Warsaw*

On October 10, 2018, Joanna Sułkowska became a laureate of the Award in the field of life sciences for her work on "looped proteins". It is worth emphasizing that the improper operation of "looped proteins" can lead to various civilization diseases, and unraveling their secrets can bring a breakthrough in the treatment of Alzheimer's disease, Parkinson's disease, or obesity.


photo: Michał Niewdana, NSC

Joanna Sułkowska specializes in molecular and theoretical biophysics. Her research is interdisciplinary – she uses physical, mathematical and biological knowledge as well as the node theory.

– *The very important thing in my work is the ability to look broadly, not to attach to a given field: analyzing the structure of proteins not only with the physicist's eye, but also taking into account the perspective of a biologist and chemist. You have to imagine things that cannot be seen with the naked eye, although they do exist* – says Joanna Sułkowska.

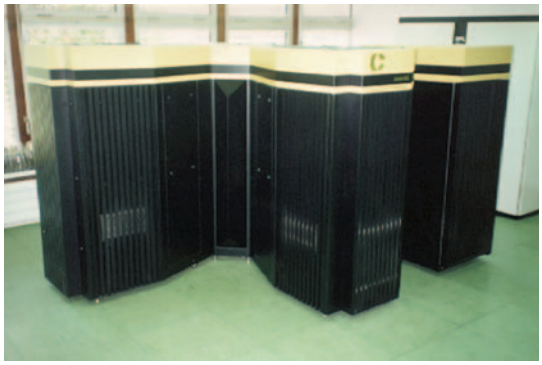

photo: Michał Niewdana, NSC

Joanna Sułkowska supports her research with calculations on the Prometheus supercomputer located at ACC Cyfronet AGH. The research is carried out within the TRMDMG3 computational grant: "Mechanism of tRNA methylation with the participation of narrowed proteins".

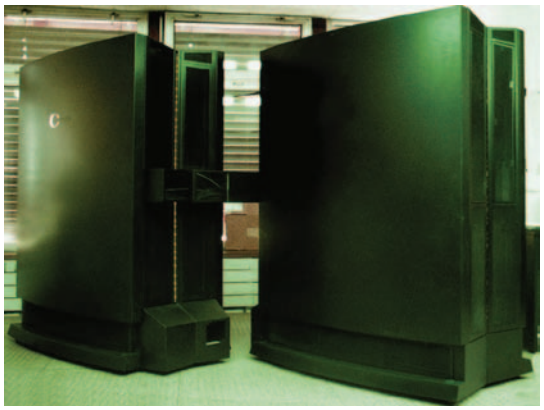
Congratulations to the winners and we wish you further success!

**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

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

MAN 10 Gbps started.

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

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

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

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

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

Total amount of installed disk space exceeds 2 PB.

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

Zeus has been placed on 81st position on the TOP500 list

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

In April, ScaleMP, a leading provider of 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

SGI Origin2000

SGI ALTIX 3700

HP Cluster Platform 3000 BL

Anniversary Medal

New Machine Hall

Prometheus supercomputer

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

Anniversary Medal has been minted

2014 The new Machine Hall is completed.

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

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

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

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

The new backup Data Center is completed.

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

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

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

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

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

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

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

EOSC-Hub and PRIMAGE projects have been launched

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

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

Cyfronet exhibition stand at the ISC'19 conference.

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

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

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

Prometheus (53,748 cores, 2.65 PFlops) ranks 288th place on the TOP500 list (the June edition).

The Prometheus supercomputer supports scientists in the fight against coronavirus.

EPOS PL +, EPOS SP, PROTEUS-RS and EUROCC 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

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

ISBN 978–83–61433–36–1