PDC Center for High Performance Computing

PDC Newsletter

**Dirk Pleiter
Director, PDC**

The PDC Newsletter is published by PDC at EECS, KTH.

PDC operates leading-edge, high-performance computers as easily-accessible national resources. These resources are primarily available for Swedish academic research and education. PDC, which is hosted by EECS, KTH, is one of the six centres in the Swedish National Infrastructure for Computing (SNIC).
Editor: Dirk Pleiter
Editorial staff: Genet Edmondson
Graphic Design: Maria Malmqvist
E-mail: support@pdc.kth.se

ISSN 1401-9671

Cover

The new HPE Cray EX system at PDC is called Dardel in honour of the Swedish author, Tora Dardel, and her first husband, Nils Dardel, a famous Swedish artist. The panels show a portrait of Tora painted by Nils, and one of Nils' most famous works, "The Dying Dandy".

How to Contact PDC

Visiting address:
Teknikringen 14, "Plan 4",
KTH, Stockholm
Mail: PDC, KTH,
SE-100 44 Stockholm, Sweden
E-mail: support@pdc.kth.se
www: http://www.pdc.kth.se
Phone: +46 8 790 7800

Editorial

This newsletter is arriving later in the summer than usual, but for a very good reason: we wanted to provide you with information about the installation of PDC's new flagship system, Dardel (see *"First Phase of Dardel Installed"*). The original plan was for Dardel to be installed in spring this year, but that had to be postponed to late in the summer for reasons beyond PDC's control.

The new system is one of several significant changes that have happened at PDC this year, which include me starting as the new director of PDC at the beginning of the year. Thanks to strong support from the other members of PDC's leadership team – namely Gert Svensson, Henric Zazzi and Lars Malinowsky, as well as PDC's interim director Patrick Norman – this change was organised as a smooth transition. Many thanks to everyone who helped with the transition and, in particular, to Patrick for making himself available as interim director. Patrick will remain closely linked to PDC through joint efforts on modern approaches to scientific software design (see *"VeloxChem: Quantum Chemistry from Laptop to HPC"*).

A third change, which will be significant for PDC, is still ahead of us: based on the feedback from an international review panel, the Swedish Research Council (VR) decided to initiate a process to establish a new organization for high-performance computing (HPC) in Sweden. This will be a good opportunity to facilitate the necessary transition of HPC resource providers (like PDC) to becoming e-infrastructure service providers serving research in Sweden while being part of the EuroHPC ecosystem, which is becoming reality as you will see in this newsletter.

In this newsletter we focus, however, on today's changes, with the arrival of Dardel being the most significant one. This will be a big step for PDC as the system comes with a lot of brand-new technology. This concerns, in particular, phase two of the system, which will be installed in autumn and facilitate a very significant boost in compute performance. While phase one of Dardel is based on nodes with very modern, but in some sense standard multi-core processors, phase two will be based on nodes comprising graphical processing units (GPUs) that act as compute accelerators. Using these GPUs requires applications to be ported using suitable programming models. The use of GPUs for scientific computing is not new, but now the choice of GPU architectures that are suitable for HPC is broadening and, therefore, portability needs to be considered. While Dardel – as well as the EuroHPC

pre-exascale system LUMI, which is currently being installed in Finland – use GPUs from AMD, other systems will continue to use products from NVIDIA or plan to use upcoming GPUs from Intel. This newsletter provides you with a primer on GPU computing (see *"Preparing for GPU Computing on the Dardel and LUMI Systems"*) as well as information on pilot projects for LUMI (see *"LUMI: Successful Swedish Pilot Projects"*) and training opportunities provided by our colleagues at the EuroCC National Competence Centre Sweden (see *"Training at ENCCS"*).

To support researchers to exploit such new hardware capabilities, PDC will continue to strongly engage with research communities, for instance, through Centres of Excellence (CoE) that are funded by the European Commission with the goal of getting researchers (and the code they use!) prepared for future exascale computing. The successful work in the BioExcel CoE, which is coordinated by PDC, continues (see *"BioExcel News"*) and two new CoEs, namely TREX and PerMedCoE, allow PDC to work closely with researchers who use quantum Monte Carlo methods for research in quantum chemistry and condensed matter physics, as well as researchers working in the area of personalised medicine (see *"TREX: Centre of Excellence for Quantum Monte Carlo Applications"* and *"PerMedCoE: Exascale-Ready Cell-Level Simulations"*). PDC continues to engage with the brain research community, and we are proud of being part of the EBRAINS research infrastructure that recently succeeded in being listed on the road map of the prestigious European Strategy Forum on Research Infrastructures (ESFRI) (see *"Performing Multi-Scale Simulations of the Brain Using MUSIC"*).

We are furthermore intensifying our efforts to make the resources available at PDC easier to use for the broader community of computational and data scientists. This involves simplifying authentication to the HPC services without compromising on security (see *"PDC Portal for Improved Login"*) as well as improved support of small-scale use of, and interactive access to, the new Dardel system. Finally, in the next newsletter, we plan to report on a broadened service portfolio after adding a private cloud instance, which can be made more openly accessible than the HPC-based services.

With this overview, I hope to stimulate your continued interest in the PDC newsletter. Let me close by thanking the PDC team for their hard work on preparing for Dardel and preparing phase one of the system to become generally available in October. With the next newsletter, we will provide you with the first results obtained on this exciting new system!

Dirk Pleiter, Director PDC

Staff Focus

Jacob Wahlgren

Jacob Wahlgren has recently joined PDC as a support technician. He helps students and researchers to use PDC's systems. Jacob's main responsibilities are answering questions sent through the PDC support form or email, and keeping documentation up to date. In addition, he assists in the task of preparing Dardel for wider use by compiling and installing software that researchers require.

Jacob is working part-time at PDC while pursuing a master's degree at KTH. He is studying computer science with a specialisation in scientific computing. Jacob's background is in software development - he previously completed a bachelor's degree in computer engineering at KTH. Jacob found his way to scientific computing through a desire to work interdisciplinarily and in an area where performance is important. Jacob is considering applying for a PhD student position after he finishes his Master's degree next year.

In his free time, Jacob tries to avoid computers by reading dead tree novels, and is still celebrating Hammarby's victory in the Swedish football cup. He also owns several instruments he barely has time to play.

First Phase of Dardel Installed

Gert Svensson, PDC

As announced in late February (see *"Press Release: New supercomputer coming to KTH!"*), Hewlett Packard Enterprise (HPE) was awarded the contract to supply a new Swedish National Infrastructure for Computing (SNIC) system at PDC. The new system, which is called Dardel, is an HPE Cray pre-exascale supercomputer that is being installed in two phases: the first phase of the system is based on central processing units (CPUs), and the second phase will be based on graphics processing units (GPUs).

Current Status

Installation of the first (CPU-based) phase of PDC's new Dardel flagship system started on the 2nd of August 2021. As with previous Cray installations at PDC, the installation of Dardel was well planned, and the system was in place in the PDC computer hall and ready to be started on the 5th of August. The photos at the end of this article show the parts of the system arriving at KTH and then being physically installed in the PDC computer hall.

According to the contract with HPE, this first phase of Dardel should have been delivered and installed at PDC in April this year. However, the installation had to be delayed due to external factors. At the time of writing (late August), the Dardel system is undergoing the first part of the acceptance testing process. When that is completed, PDC staff will spend two weeks integrating the

---

Save the Date

BioExcel Conference 18-21 October 2021
EMBO Workshop:
Advances and Challenges in Biomolecular Simulations

For more details and to register, see https://bioexcel.eu/events/embo-workshop-advances-and-challenges-in-biomolecular-simulations.

ENCCS/NVIDIA GPU Hackathon
6 & 13-15 December 2021

GPU hackathons offer a unique opportunity for researchers and software engineers to accelerate and optimise their applications on graphics processing units (GPUs). Teams of researchers are paired with experienced GPU mentors to learn and apply accelerated and parallel computing skills to the teams' own projects. Both current and prospective users of large hybrid CPU/GPU high-performance supercomputers who develop applications that could benefit from GPU acceleration are encouraged to participate! For more details and to register, see https://www.gpuhackathons.org/event/enccs-gpu-hackathon-2021.

Below: This table shows the number and types of nodes that will make up the first phase of Dardel. The coloured boxes indicate the nodes that have already been installed.

Name of nodes	Memory	Number of nodes					
		SNIC initial	Industry	Scania	SU Dept. of Astronomy	SNIC additional	Total
Thin	256 GB	488	36				524
Large	512 GB	20		236	12	48	316
Huge	1 TB	8					8
Giant	2 TB	2				8	10
Total		518	36	236	12	56	858

new system into the PDC environment. This will involve tasks like connecting the system to PDC's network, setting up user authentication, and installing application software. After that, there will be a one-month test period with selected test users trialling the system and trying to push its boundaries. If problems arise, there may need to be additional test periods after the problems are mitigated. Once a successful test period has been completed, all researchers will gradually be moved over from Beskow and Tegner to Dardel. We expect that all PDC users will be moved over during October at the latest, and then it will be possible to decommission Beskow and Tegner.

CPU Nodes

The original contract for the new system specified that phase one of Dardel would contain 518 computational nodes with dual AMD 7742 CPUs with 64 cores each. Using options in the initial agreement, PDC has been able to significantly expand the system with additional nodes as follows.

- An extra partition with 36 nodes for industrial and business collaboration at PDC is being included in phase one.
- Phase one of Dardel will also have an additional partition consisting of 236 computational nodes to be used in the research collaboration between KTH and the heavy vehicle manufacturer Scania.
- Furthermore, SNIC has decided to invest in 56 additional nodes for academic research.
- The Department of Astronomy at Stockholm University (SU) has also invested in 12 nodes.

The Scania partition, the SU partition and the additional SNIC nodes have not been installed at

the time of writing. Details of all the nodes in phase one can be found in the table above; the nodes that have been installed to date are indicated.

GPU Nodes

The second phase of the Dardel system is a small but powerful partition consisting of 56 GPU nodes, each with four future AMD Instinct GPUs. The GPU nodes are expected to be delivered at the end of this year and will increase the system's total capacity by more than a factor of two.

Updates to the PDC System Environment

In conjunction with the installation of Dardel, some updates in the previous PDC system environment are planned. For example, there will soon be a choice of login methods: using either a Kerberos-based approach, or a Secure Shell (SSH) login method relying on key pairs. This is discussed more in the article *"PDC Portal for Improved Login"*.

The PDC system environment has been utilising two systems for storing data files: AFS and Lustre. The AFS system has primarily been used for user's home directories; small amounts of data could be stored there in the longer term. PDC's Lustre system is a parallel file storage system that has mainly been used to provide extremely fast access to large amounts of data for running simulations. However, that situation is about to change. Existing home directories in the AFS file system will be replaced by home directories in the Lustre system. Note that the plan is for the AFS file system to still be available from the Dardel file transfer nodes, so researchers will be able to transfer any data they have stored in the AFS file system to Lustre.

Preparing the PDC computer hall and installing Dardel

The photos to the left show some of the preparatory work in the PDC computer hall earlier this year. The photos to the right and on the next page show scenes from the week of 2-5 August 2021 when Dardel was delivered and installed at PDC.

For more photos of these events, plus captions that explain what is happening in these photos, see https://www.pdc.kth.se/hpc-services/computing-systems/photo-timeline.
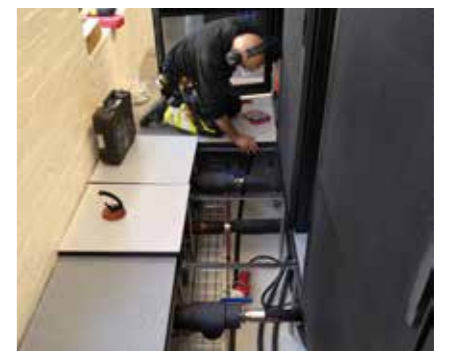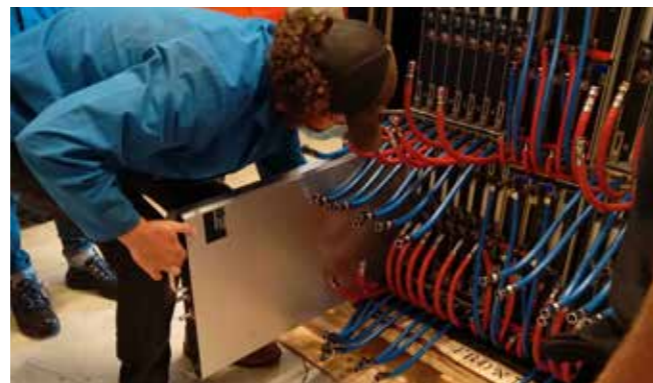
Preparing for GPU Computing on the Dardel and LUMI Systems

Peter Larsson, PDC/LUMI User Support Team

Did you know that you can get started on programming for AMD GPUs right now using the ROCm software stack on NVIDIA GPUs?

Early next year, phase two of PDC's new system, Dardel, and the European pre-exascale LUMI system will come online with their powerful new graphics process units (GPUs) from AMD. The exact technical specifications of these GPUs are still under wraps, but it is expected that they will offer better performance than the GPUs available today. This will be a big shift in the Swedish computing landscape as researchers have not previously had access to large GPU computing resources at the national level. Starting next year, most of the Swedish compute capacity will be GPU-based with NVIDIA GPUs in the Alvis and Berzelius clusters, which are located at the Chalmers Centre for Computational Science and Engineering (C3SE) and the National Supercomputer Centre at Linköping University (NSC) respectively, plus AMD GPUs in Dardel and LUMI. This article is about what you can do now to start preparing to take advantage of the AMD GPUs when they come online next year.

If you are a researcher using software that others have written, you should first check the documentation for clues about GPU usage. Does the software support GPUs at all? If so, can you find anything about the way that software has been programmed to use GPUs? The keywords to look for are "CUDA", "HIP", "OpenCL", "OpenMP offloading", "OpenACC" and "SYCL". A checklist is presented in the table below. The most important thing is that CUDA, which is a propriety framework from NVIDIA for developing applications for their own GPUs, will not be available on LUMI and Dardel. This means that any programs using it will need to be updated in order to work with the AMD GPUs. The same applies to many of the popular libraries for NVIDIA GPUs, such as cuBLAS, cuFFT, and cuDNN. Instead, for AMD GPUs, the corresponding language for GPU programming is called HIP. If you see that mentioned, then it is great news but, unfortunately, only a few software packages have been modified to use HIP so far. Furthermore, if OpenCL is mentioned, then the software should also work on AMD GPUs, even though the performance may not be optimal. For high-level GPU programming, there are three popular models: OpenACC, OpenMP and SYCL. If the software uses OpenACC for GPU computing, then it might work eventually, but will not work right now (which is discussed further on the next page). If the software already uses OpenMP offloading for GPU computing, then it should work well on LUMI and Dardel, as this is the recommended high-level way to write GPU code for these systems. SYCL is gaining popularity, as it is backed by Intel for their upcoming GPUs, and it is possible that SYCL will work well on AMD GPUs in the future, but right now it is an experimental project (hipSYCL at Heidelberg University).

If you are a software developer writing your own low-level GPU code, you should investigate the ROCm framework from AMD and the HIP language extension for C/C++. ROCm and HIP are intended to be a full replacement of the CUDA stack with the added advantage that code written in HIP can be compiled for both AMD and NVIDIA GPUs! This allows you (in theory) to have a single codebase that supports both. It is important to realise that this means that you can start porting your code to HIP today and test it on NVIDIA GPUs, even if you do not have access to new AMD GPUs that support ROCm. This is the recommended approach as supplies of new GPUs are extremely tight due to high consumer demand and supply chain problems as a result of the coronavirus pandemic. The HIP language is designed to be very similar to CUDA. In some cases, it is possible to automatically translate CUDA code to HIP code using the "hipify" tool in ROCm. The hipify tool will even translate CUDA library calls to the corresponding AMD library calls but, in general, some code modifications are needed to get the best performance. HIP code can then be compiled either with Clang (which is included with ROCm) or Cray's compiler (which only has experimental support for now). Courses in HIP are already starting to appear; there was one at the CSC – IT Center for Science (CSC) in Finland in February 2021 and another in Sweden in April 2021 that was jointly arranged by the EuroCC National Competence Centre Sweden (ENCCS) and CSC, and there will be more to come.

If you are programming for GPUs using a high-level approach such as OpenACC, OpenMP device offloading, or more recently SYCL, you need to investigate the support for AMD GPUs in your toolchain. OpenMP offloading is likely the best choice as it is the only cross-platform framework supported by the big three GPUs vendors (NVIDIA, AMD, and Intel). SYCL could be a good choice in the future when there is a backend for AMD GPUs. The most problematic is likely to be OpenACC. In practice, OpenACC has been heavily centred around NVIDIA products, which made it unclear how well it would be supported on AMD GPUs. There is hope that the "Clacc" project will solve this problem. It is a project sponsored by the US Exascale Computing Project. It aims to develop an OpenACC compiler for Clang that will effectively translate OpenACC to OpenMP. This approach should allow OpenACC code to run on AMD GPUs in the future. Recently, Cray also announced that they will provide full support for OpenACC in their compilers for both C/C++ and Fortran, likely based on the Clacc project. Initially, there will be support at the OpenACC 2.7 level starting next year and eventually at the OpenACC 3.0 level. This should help to make some major software packages (like VASP) use GPUs on LUMI and Dardel.

Many applications rely on subroutines from optimised numerical libraries, such as BLAS and FFTW, to get good performance on CPUs. Several of these libraries are also available for NVIDIA GPUs, which has helped the uptake of GPU computing. A similar ecosystem is currently under development for AMD GPUs, but many of the well-known GPU libraries have actually already been ported (see the table below). It is a good idea to continue this approach with AMD GPUs

Programming model	Nvidia GPUs	AMD GPUs	Intel GPUs
CUDA	yes	no, but HIPIFY tools may help with conversion	no
HIP	yes	yes, best performance	?
OpenCL	yes, but likely lower performance	yes, medium performance	yes
OpenMP (offloading)	yes	yes	yes
OpenACC	yes (Cray/Nvidia compilers)	yes (in the future through Clac?)	yes (in the future through Clac?)
SYCL	yes (Codeplay's ComputeCpp product)	experimental support (HipSYCL)	yes (Intel OneAPI)

Above: Overview of GPU programming models and their support on different kinds of GPUs

Description	CUDA library	HIP library	ROCm backend
Basic linear algebra like matrix-matrix multiplication	cuBLAS	hipBLAS	rocBLAS
Fast Fourier transforms	cuFFT	hipFFT	rocFFT
Linear algebra (subset of LAPACK)	cuSOLVER	hipSOLVER	rocSOLVER
Basic linear algebra for sparse matrices	cuSPARSE	hipSPARSE	rocSPARSE
Parallel algorithms like scan and reduce	CUB	hipCUB	rocPRIM
Random number generation	cuRAND	hipRAND	rocRAND
GPU to GPU communication	NCCL ("Nickel")	n/a	RCCL ("Rickle")
Neural network operations	cuDNN	hipDDN	miOpen

Above: Translation table from CUDA libraries to the corresponding HIP libraries
The libraries are available under https://github.com/ROCmSoftwarePlatform.

and rely on libraries when you can. This way, you will automatically get most of the performance benefits with future AMD GPUs without changing your own code, as the libraries are likely to be updated to support the latest features. Generally, the ROCm libraries come in two flavours: there is a "hip-xxx" version which can run on both AMD and NVIDIA hardware, and a "roc-xxx" version which only runs on AMD GPUs. The HIP version is just a thin library of wrappers that call the best underlying library depending on what GPU the code is compiled for. For example, on an NVIDIA system, hipBLAS will call cuBLAS, but on an AMD system, hipBLAS will call rocBLAS instead.

Finally, what software is already available for AMD GPUs? In general, the fields of deep learning and molecular dynamics have the most software packages which are either OpenCL-based or have early HIP ports. The deep learning frameworks Tensorflow and Pytorch can already run on AMD GPUs right now, and the performance looks promising. For molecular dynamics, several of the big molecular dynamics packages have support or are being ported: Gromacs, LAMMPS, NAMD, and Amber. In the materials science field, there is some support in CP2K and Sirius through the underlying DBCSR sparse matrix-matrix multiplication library, but it is not complete. There has been no official statement on VASP, but their GPU functionality is being moved from CUDA to OpenACC, which could then work in the future when there is OpenACC support for AMD GPUs. In the weather and climate field, there are ongoing efforts to support porting the ICON modelling framework and, in the area of computational fluid dynamics, the NEK5000 code is also being prepared. In order to find AMD GPU-enabled software, it might be worth looking into projects associated with US National Laboratories and the Exascale Computing Project (ECP) as they are betting heavily on the AMD-GPUs for their upcoming Frontier and El Capitan supercomputers.

If you have questions about preparing your research code for GPU computing, you are welcome to contact Peter Larsson (ypetla@kth.se).

TREX: Centre of Excellence for Quantum Monte Carlo Applications

Dirk Pleiter, PDC

Almost a century after Erwin Schrödinger postulated the equation bearing his name, it remains a big challenge to solve this equation in the many-body case. However, this is needed in order to describe quantum mechanical electron problems that occur in the areas of quantum chemistry and condensed matter physics. Therefore, solving this equation is of practical relevance in applied research, for example, for designing novel materials with specific properties. Thanks to numerical methods, computers can help to address this challenge. But the complexity of the problem is extremely high. To cope with the high complexity, either approximate deterministic methods or stochastic techniques are being used. Quantum Monte Carlo (QMC) methods belong to the latter. With the help of QMC methods, properties of molecular and extended systems at the nanoscale can be computed with high accuracy. These types of simulations are highly computationally intensive and require supercomputers like those at PDC.

QMC methods have in common that they use Monte Carlo algorithms to evaluate integrals over a very large number of dimensions. Within such algorithms, numerical results are obtained by computing a very large number of random samples that are created with a defined statistical distribution. While this approach makes it possible to avoid the use of approximations with difficult-to-control systematic errors, the disadvantages of Monte Carlo techniques are their slow convergence properties and the need for large amounts of computational

resources to keep statistical errors small. Fortunately, QMC methods are embarrassingly parallel and therefore scale very well on the largest available supercomputers. As a consequence, applications based on QMC methods are likely to become killer applications for upcoming exascale systems and are meanwhile playing an important role in different initiatives to prepare for the upcoming exascale era.

In Europe, such an initiative is driven by the European Commission (EC) through its H2020 program and recently through the EuroHPC Joint Undertaking (https://eurohpc-ju.europa.eu). Through this initiative, various Centres of Excellence (CoEs) have been created, several of them with the involvement of PDC. Most recently, PDC became part of the TREX CoE, where the acronym stands for "Targeting Real chemical accuracy at the EXascale". It is coordinated by the University of Twente in the Netherlands and also involves partners from Austria, France, Germany, Italy, Poland, Slovakia, and the UK. Most of the partners are developers of QMC applications, while partners like PDC contribute their expertise in high-performance computing.

The most important goal of TREX is to prepare applications for future supercomputers, which is challenging for a variety of reasons. Supercomputer architectures have become significantly more complex and often also heterogeneous. For example, supercomputers like the new Dardel system at PDC leverage not only the compute capabilities of processors but also take advantage of compute accelerators - like graphics processing units (GPUs) - which deliver most of the compute

performance. Additionally, the technologies in modern supercomputers have become more diverse. To make investments in complex scientific software sustainable, not only code portability but also performance portability have become key concerns. An application is considered performance-portable if it is able to run efficiently on different architectures. Achieving the aforementioned goal of TREX is challenging - particularly as many of the codes that are currently in use started out being developed decades ago without the developers having any knowledge of today's computer architectures and without the use of more recent software engineering concepts that make it easier to achieve portability and performance portability.

TREX brings together the developers of various QMC applications that are developed in Europe. This allows for a strategy where performance-critical routines, which several of these applications have in common, can be identified and moved into a library called QMCkl. This new library will not only be developed jointly but will also be developed with performance portability in mind. The ambition of TREX is to deliver applications that can run efficiently on a diverse set of future supercomputers, where the architectures may not even be known yet. These range from supercomputers like Dardel at PDC (with its new AMD GPUs) to future supercomputers with ARM-based processors such as those currently being developed by the European Processor Initiative.

For more information about the TREX CoE, visit https://trex-coe.eu or contact Dirk Pleiter (pleiter@kth.se).
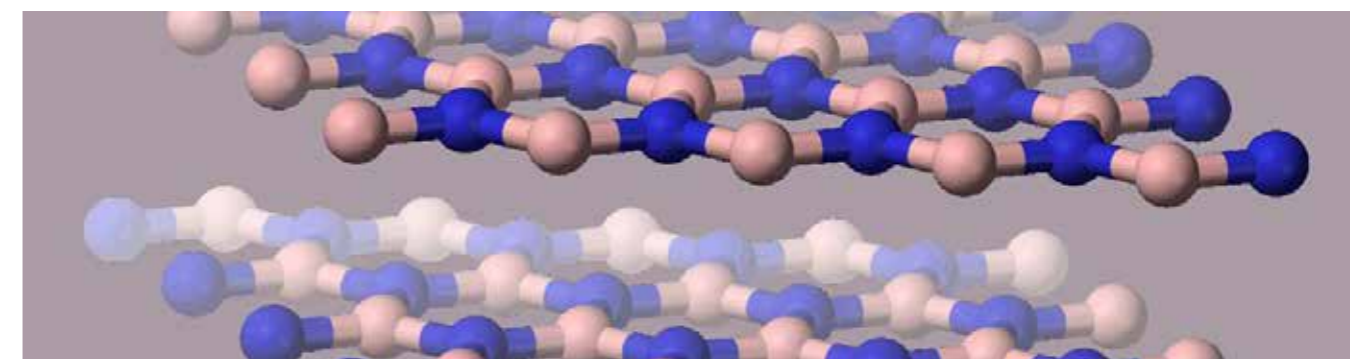
Above: TREX focuses, amongst other things, on layered materials like graphene and borophene that are promising for next-generation electronic devices. Predicting the structure and functionality of such compounds requires accurate treatment of van der Waals forces, which represents a major computational challenge that is being addressed with QMC methods.

PerMedCoE: Exascale-Ready Cell-Level Simulations

Rossen Apostolov, PDC

PerMedCoE is the high-performance computing (HPC) and exascale Centre of Excellence (CoE) for personalised medicine in Europe. It was launched on the 7th of October 2020, and is coordinated by the Barcelona Supercomputer Center (BSC) and funded by the European Commission.

Personalised medicine involves tailoring medical treatments to individual patients, rather than assuming that all patients displaying the same symptoms should receive the same treatment. The concept of personalised medicine is based on the idea that different patients may need differing treatments depending on factors such as their genetic background, previous medical history, age, diet and so forth. It is obviously a complex problem to establish models that take the large range of potential variables into account and to analyse existing medical data looking for patterns associated with given factors that may influence the most suitable treatment for any given individual. Consequently, the next generation of exascale supercomputers will be necessary tools for treating diseases on an individual level and making significant steps forward in the field of personalised medicine.

PerMedCoE will provide an efficient and sustainable infrastructure to support the development of personalised medicine by showcasing HPC/exascale-upgraded use cases, such as translating the consequences of single-cell omics information into actionable molecular disease models. (Omics refers to the biological fields ending in "omics", such as genomics and proteomics.) One of the most significant contributions of PerMedCoE will be to scale-up cell-level simulations to the HPC/exascale level, which will bridge the gap between the molecular- and organ-level simulations provided by other European Centres of Excellence, such as CompBioMed (the CoE focused on the use and development of computational methods for biomedical applications) and BioExcel (the CoE for computational biomolecular research), and thus contribute to the European Personalised Medicine Roadmap.

Personalised medicine opens unexplored frontiers for treating diseases at the individual level by combining clinical and omics information. However, the performance of current simulation software is still insufficient to tackle medical problems such as tumour evolution and response to treatment at the single-cell level or patient-specific treatments. For this reason, PerMedCoE is pursuing the following main objectives:

- to optimise key software for cell-level simulations to the new pre-exascale platforms to contribute to the European Personalised Medicine Roadmap,
- to integrate personalised medicine into the new European HPC/exascale ecosystem, by offering access to HPC/exascale-adapted and optimised software,
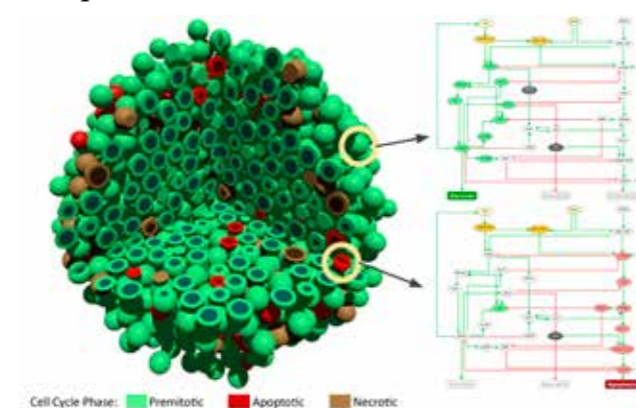
Above: This image (produced by Miguel Ponce de León from PerMedCoE) illustrates one example of the framework that makes multiscale simulations possible. PerMedCoE will enable researchers to scale-up multiscale simulations which combine Boolean models that capture intracellular behaviours and agent-based models for studying population dynamics.

- to design and complete a comprehensive set of personalised medicine use cases to drive the development of cell-level simulations, and
- to build the basis for its sustainability by coordinating personalised medicine and HPC communities, and reaching out to industrial and academic end-users with use cases, training, expertise and best practices.

The centre will become the entry point to exascale-ready cell-level simulation software, able to transform personal omics data into actionable mechanistic models of medical relevance, supporting developers and end-users with know-how and best practices. It will also connect simulation software developers with HPC, high-throughput computing (HTC) and high-performance data analytics (HPDA) experts at the POP and HiDALGO Centres of Excellence, and work with other biomedical consortia, such as ELIXIR and LifeTime, as well as connecting pre-exascale infrastructures hosted by supercomputing centres (such as BSC and the CSC – IT Center for Science in Finland).

To test its core applications and help prepare for exascale computing, PerMedCoE has designed five biologically relevant use cases to serve as pilot projects. These use cases were selected to reflect a broad range of computationally-demanding real-life biomedical scenarios in which to use cell-level models. Of the five cases, the pilot projects concerning the study of drug synergies and the study of COVID-19 have been prioritised. More information about these (and the other use cases) is available at https://permedcoe.eu/use-cases-on-drug-synergies-and-covid-19.

PDC is participating in PerMedCoE as a strong link to BioExcel (which is coordinated by PDC) as both CoEs have complementary activities. PDC will provide scientific expertise and assistance with the usage of GROMACS and molecular dynamics simulations in general. In addition, PDC will advise PerMedCoE on how to run a successful CoE and on establishing a sustainable research infrastructure. For further information on PerMedCoE, see https://permedcoe.eu or contact Rossen Apostolov (rossen@kth.se).

Training at ENCCS

Thor Wikfeldt, ENCCS

Immediately after the establishment of the EuroCC National Competence Centre Sweden (ENCCS) in September last year, the ENCCS team started drafting an ambitious plan for delivering training events on current topics in high-performance computing (HPC), high-performance data analytics (HPDA) and artificial intelligence (AI). ENCCS' raison d'être is to develop competence and knowledge in Swedish public sector organizations, industry and academia, and to enable researchers to take advantage of forthcoming (pre-)exascale EuroHPC resources as well as modern AI methodologies.

We are now happy to report that over 500 people have signed up to our training workshops and events so far, and that number is rising quickly as we are now delivering multiple events every month. In part, this is made possible by a number of excellent collaborations that have been established with external training providers, including CSC, different Swedish National Infrastructure for Computing (SNIC) centres, RISE, CodeRefinery, HiDALGO, NVIDIA, and Intel, amongst others. If you would like to browse through training material from earlier events, please visit https://enccs.se/training-resources where you will find links to public resources that, in most cases, are also suitable for self-learning at your own pace!

To be sure to catch any future events that could be of interest to you, you can visit our events calendar at https://enccs.se/events, subscribe to our newsletter (https://enccs.se/newsletter) or follow the project on Twitter (https://twitter.com/EuroCC_Sweden).

Above: Some of the topics in the ENCCS training resources

Human Brain Project

## Performing Multi-Scale Simulations of the Brain Using MUSIC

Mikael Djurfeldt, PDC

The brain can be investigated by making mathematical models at different scales. For example, one could consider how different regions of the brain (such as the cortex and the basal ganglia) interact with each other, or investigate the electro-chemical interactions between individual neurons. When building a mathematical model of the brain, a single scale is usually considered, and similarly most brain simulation software is specialised for just one scale of the brain. For example, the neuronal network simulator NEST is specialised for simulating networks of point neuron models, while the simulator NEURON is specialised for networks of compartmental neuron models where distinct ionic currents, as well as their spatial distribution, are taken into account. However, there are times when brain modelling is treated as a multi-scale problem where different models cover different scales. In such situations, it can be desirable to use different simulators or even neuromorphic systems. This raises the question of how to effectively combine the different models and associated software effectively.

MUSIC is one solution to this problem. It is a type of coupling software, developed at PDC, that can be used for applications in the computational neuroscience domain. MUSIC makes it possible to co-simulate a single brain model covering multiple scales by coupling different simulators, such as NEST and NEURON, as well as neuromorphic hardware and virtual environments.

The development of MUSIC has been supported by the Human Brain Project (HBP), a flagship project funded by the European Commission. As such, MUSIC is part of an effort to create a digital research infrastructure, which is called EBRAINS. Earlier this year, EBRAINS was accepted

Below: Multi-scale modelling of the brain can be performed by using the MUSIC software to couple packages for simulating different scales of the brain, such as NEST (for simulating networks of point neuron models) and NEURON (for simulating networks of compartmental neuron models). The instances of MUSIC communicate using the message passing interface MPI.
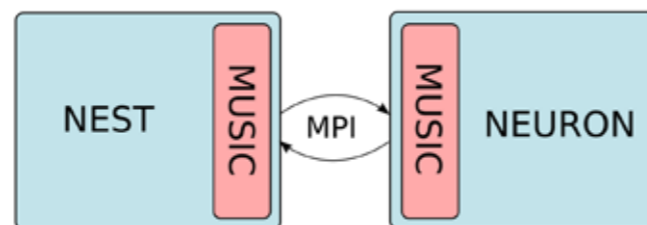

as a European Strategy Forum on Research Infrastructures (ESFRI) project (see https://www.esfri.eu/latest-esfri-news/new-ris-roadmap-2021).

The EBRAINS infrastructure offers a collection of services to the neuroscience community through a set of national nodes, and planning for the creation of a Swedish EBRAINS node is currently underway. In the future, the MUSIC software may be made available to brain researchers and software developers through EBRAINS in conjunction with a range of support services (such as training researchers to use MUSIC for multi-scale brain simulations, training application developers so their software can easily be incorporated into MUSIC, and implementing new features in MUSIC to further the possibilities of multi-scale brain simulations).

At this stage, the next steps in the development of MUSIC are focused on: extending the MUSIC-neuromorphic hardware interface, providing support for unified descriptions of MUSIC-enabled multi-scale models in the popular model specification language PyNN, and improving the MUSIC documentation and adding better support for debugging.

If you are interested in using MUSIC for your brain-related research or developing applications to work with MUSIC, please contact Mikael Djurfeldt (mdj@pdc.kth.se) for further information.
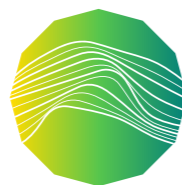
Great Success of Swedish Researchers in the PRACE DECI17 HPC Systems Access Call

Lilit Axner, ENCCS

Nine Swedish applications for compute resources to European high-performance computing (HPC) systems have been successfully accepted through the DECI17 call for proposals that closed on the 31st of January 2021 with a project start date on the 1st of June 2021. The projects are as follows.

- Dr. Walter Villanueva (KTH), *Large-scale Simulations of Industrial Flows (LSIF)*
- Prof. Mihai Mihaescu (KTH), *Convective Heat And mass transfer in INternally heated Molten pools (CHAIN)*
- Prof. Natalia Skorodumova (KTH), *Multiscale Modelling of defective Ceria at finite temperatures (P2021CEO2SB)*
- Prof. Andrei Ruban (KTH), *Ab Initio modelling of chemical vapor deposition for efficient computational design of new advanced coatings (AMCVD2)*
- Dr. Zilvinas Rinkevicius (KTH), *Magnetic properties of nanodiamonds: A large-scale ab initio modeling*
- Dr. Torbjörn Larsson (Faurecia Creo AB), *Accurate and Efficient CFD Predictions of Fast Re-Fueling of Hydrogen Storage Tanks (H2CFD)*
- Dr. Camilla Gustafsson (EnginZyme AB), *Engineering enzymes for increased stability (EnginZyme)*

- Prof. Biplabb Sanyal (UU), *Computational design of novel quantum states in 2D topological materials (Q2Dtopomat)*
- Prof. Henrik Grönbeck (CTH), *Multiscale modeling of nanoparticle catalysis (NANOCAT)*

ENCCS, the Swedish National Competence Centre in the HPC, HPDA (high-performance data analytics) and AI (artificial intelligence) arena, assisted four out of these nine successful applicants with the DECI17 application procedure. Two of these four applicants are Swedish small to medium-sized enterprises (SMEs), and it is wonderful to see SMEs being awarded research time on the DECI systems! European SMEs are especially welcome to apply for PRACE resources and Sweden is topping the success rate charts with the number of applications from Swedish SMEs.

Starting from the 20th of April 2021, the EuroHPC Joint Undertaking (EuroHPC JU) strengthened the European HPC ecosystem with even more powerful supercomputers which are in the pre-exascale and petascale range. These systems are now available to Swedish researchers (from academia, industry or the public sector) to be used for their research computations, and the calls for access are also open through the same PRACE-EuroHPC portal (https://prace-ri.eu/hpc-access/eurohpc-access). ENCCS welcomes all types of applicants (whether from the business, academic or public sectors) to contact us for assistance with application procedures for both PRACE and EuroHPC JU systems access calls.

Save the Date

SONC Meeting 28-29 October 2021

The next SNIC Operators' Network of Competence (SONC) meeting, which is for system administrators at SNIC centres, will be organized by PDC and held on the 28th and 29th of October 2021. More details will be made available soon.

VeloxChem: Quantum Chemistry from Laptop to HPC

Patrick Norman, Department of Theoretical Chemistry and Biology, KTH

On the 6th and 7th of May 2021, PDC, together with the EuroCC National Competence Centre Sweden (ENCCS), offered a training event on the VeloxChem program (https://veloxchem.org). Guided by a team of program developers, some 40 trainees (from master's degree students to professors) participated in the event and learned how to use the VeloxChem program package to do the following.

- Perform quantum chemical simulations of ground- and excited-state properties of large systems
- Leverage the aggregate resources of modern high-performance computing (HPC) clusters to efficiently tackle large molecular systems
- Use the Python application programming interface (API) to prototype new methods
- Design interactive computational teaching materials in Python

The workshop took place via Zoom and included topical presentations by the mentors, as well as breakout sessions where participants completed semi-prepared Jupyter notebooks on Binder and submitted multi-node jobs on supercomputers; an allocation on Beskow was made available for this purpose. A special acknowledgement goes to Xin Li from PDC and Roberto Di Remigio from ENCCS for preparing and organizing this event, which was very well received by the participants as evidenced by the feedback in the subsequent survey. All the material developed for the workshop is freely available via the ENCCS home page (https://enccs.github.io/veloxchem-workshop), so it can continue to be explored, utilised, and learned from. Anyone interested in using VeloxChem for research or educational purposes is welcome to browse the material!

More information about the VeloxChem project can be found in the cover article of the PDC Newsletter No. 2, 2019 (https://www.pdc.kth.se/publications/pdc-newsletter-articles/2019-no-2/veloxchem-1.945304), where the dual aspects of the project goals – namely being science- and education-enabling – are clearly seen to be in focus. In terms of the former aspect, the team of developers are presently hard at work to make VeloxChem ready for use on the new Dardel system at PDC, first on phase one and later also on phase two of the system. Once this point is reached, we will offer a new VeloxChem training event to give a taste of the range of opportunities for quantum molecular modelling that will be available utilising the combination of VeloxChem and Dardel.

```
In [ ]:  import veloxchem as vlx

         h2o_xyz = """
         O    0.000000000000     0.000000000000     0.000000000000
         H    0.000000000000     0.740848095288     0.582094932012
         H    0.000000000000    -0.740848095288     0.582094932012
         """

         molecule = vlx.Molecule.read_str(h2o_xyz, units='angstrom')
         basis = vlx.MolecularBasis.read(molecule, "6-31g")

         scf_drv = vlx.ScfRestrictedDriver()
         scf_drv.compute(molecule, basis)
```

Program reference

Rinkevicius, Z.; Li, X.; Vahtras, O.; Ahmadzadeh, K.; Brand, M.; Ringholm, M.; List, N. H.; Scheurer, M.; Scott, M.; Dreuw, A.; Norman, P. VeloxChem: A Python-driven Density-functional Theory Program for Spectroscopy Simulations in High-performance Computing Environments. WIREs Comput. Mol. Sci. 2020, 10 (5), 1–14. https://doi.org/10.1002/wcms.1457.

Above: Example of VeloxChem in Jupyter notebooks

LUMI

LUMI: Successful Swedish Pilot Projects

Lilit Axner, ENCCS

LUMI (https://www.lumi-supercomputer.eu), one of the EuroHPC (https://eurohpc-ju.europa.eu) pre-exascale supercomputers, is located at the CSC – IT Center for Science (CSC) data centre in Kajaani, Finland. The supercomputer is hosted by the LUMI (Large Unified Modern Infrastructure) consortium, which includes ten European countries: Belgium, the Czech Republic, Denmark, Estonia, Finland, Iceland, Norway, Poland, Sweden, and Switzerland. By bringing together their unique high-performance computing (HPC) expertise and experience, these countries will together provide added value for the whole of Europe in relation to HPC and the LUMI pre-exascale resource.

The LUMI system is being installed in two phases: the first part is based on central processing units (CPUs), and the second part of the system will feature graphical processing units (GPUs). The plan is for the CPU part of LUMI to be operational for EU users in the third quarter of 2021, while the GPU part of LUMI will be operational in the first quarter of 2022. However, before the system can be considered operational, it will need to pass an extended list of tests, and that is the case for both the CPU and GPU parts (or phases) of the system. For testing purposes, each LUMI consortium country has selected up to two projects for each phase. The selected pilot projects will be run on the relevant phase of the system to test that phase to its limits.

The call for the Swedish pilot projects was organised by the Swedish Research Council (VR) and the Swedish National Infrastructure for Computing (SNIC), and it was open to researchers in Swedish higher education and research institutions or in Swedish enterprises. In order that the pilot projects could be used to test the system thoroughly, the selected projects were required to have specific profiles.

Pilot projects to run on phase one of the system needed to involve data-intensive computing, high-throughput computing, and high-performance data analytics (CPU-only) workloads to generate stress, especially on the storage systems.

For phase two, projects had to utilise highly scalable GPU applications to keep the GPU partition under full load and demonstrate its scalability.

In addition, VR and SNIC required that the completed projects would be able to demonstrate scientific significance, as well as having the potential to provide breakthroughs in the relevant research area, and have a positive impact on society.

Ten Swedish applications were received for these two-phased pilot calls. Out of those applications, four projects were selected to be run as pilots on LUMI, with two projects being run in each testing phase. The four pilots representing the Swedish part of the consortium are as follows.

- *The baryon cycle in colliding galaxies*, Prof. Oscar Agertz, Lund University
- *Turbulence data generation on Boeing hump*, Prof. Philipp Schlatter, KTH Royal Institute of Technology
- *Next generation global climate models*, Prof. Thorsten Mauritsen, Stockholm University
- *Lipid modulation of ion channel gating - showcasing the GROMACS AMD GPU port*, Prof. Erik Lindahl, KTH Royal Institute of Technology

For all four projects, the EuroCC National Competence Centre Sweden (ENCCS) has been assisting the principal investigators with the process. Additionally, ENCCS is committing substantial effort in three out of the four projects to support porting the necessary software to LUMI.

Researchers from academia or business/industry who are interested in applying to use LUMI after the pilot period can contact ENCCS for support with preparing applications (see https://enccs.se/proposal-support).

PDC Portal for Improved Login

Michaela Barth, Harald Barth & Gert Svensson, PDC

Protecting identities and data is critically important these days. When researchers from academia or business/industry log in from their personal computers or laptops to use PDC's systems for their research, they are communicating over non-secure networks. The identity, research work and data of the researchers who use PDC's systems need to be protected from security threats and attacks, as do PDC's own resources. As an example, last summer saw security breaches at various European high-performance computing (HPC) institutions. The results of such breaches can be that systems are down for weeks, or research data can be stolen or corrupted. It is therefore vitally important for users to have a secure method for logging in to PDC's systems.

A current trend being seen on a European level is to use what is known as multi-factor authentication for security purposes rather than single-factor authentication. Authentication of identities that relies on only one factor is not regarded as secure. For example, in the past, a single password was used to log in to online bank accounts, so anyone who found out the password could log in to the account. Nowadays, European banks use two-factor authentication (2FA) in accordance with the EU's Strong Customer Authentication (SCA) requirements. This means that two identifiers are needed: one could be a personal password; the other could, for example, be provided via an SMS or through a One-Time Pad (OTP) application like Google Authenticator. It is expected that similar national regulations for accessing information systems hosted by academic organizations (such as the MSBFS2020:7 security measures in relation to information systems for Swedish authorities, https://www.msb.se/sv/regler/gallande-regler/krisberedskap-och-informationssakerhet/msbfs-20207) will be extended to encompass all users.

PDC is in the process of developing a secure and reliable login approach which will be used for logging in to the new Dardel system, and which would in principle scale up to work for accessing all HPC systems in Sweden available through the Swedish National Infrastructure for Computing (SNIC). Many of the academic and business researchers who use HPC systems in Sweden for their research actually work on several Tier-1 systems within Sweden (through SNIC) and may also use Tier-0 systems in Europe (through PRACE or EuroHPC). With the example of the new Dardel login approach, PDC hopes to influence the debate in Sweden concerning a login method that could be implemented on all the SNIC systems over time, and that would ideally make it easy for users who want to migrate their research to a Tier-0 system outside of Sweden (such as the EuroHPC system LUMI).

Of course, the new approach should be easy-to-use, so preferably any cumbersome aspects of the authentication process should only be done when initially setting up communications between the user and the system. Once such a "session" is established, it could be trusted for a certain validity period and thus researchers would not have to go through a full authentication process every time they (re-)connected to the system.

PDC's proposal is to use the SNIC User and Project Repository (SUPR) to provide authentication for managing login set-up in a future-proof way. In coming years, it is likely that a European identity provider (IdP), such as GÉANT, will act as a proxy IdP for accessing HPC research systems within Europe. The PDC solution would be designed to be compatible with that potential scenario. Such an approach would have the advantage that users would only need a single (virtual) identity and ideally only one set of credentials to use various HPC systems, rather than having to identify themselves in different ways to access different research systems. The plan is that, in the first stage, a single authentication solution would be provided nationally through SUPR in Sweden, and 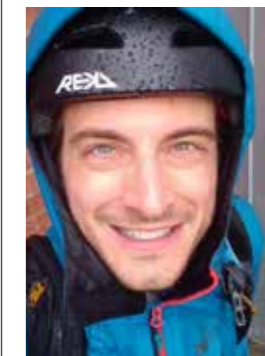then, in the longer term, that would be extended to apply across all of Europe. This approach makes sense as SNIC already has a well-established and proven authentication solution, known as SUPR 2FA, which is used for all the SNIC centre staff and for the principal investigators (PIs) using the SNIC SENS computing and storage systems for sensitive data.

The new login approach at PDC also needs to take into account the fact that many researchers log in repeatedly on days when they are using PDC's systems. To make it easier for these users – and to keep the process of automating jobs as simple as possible – multi-factor authentication would not be required for multiple connections by a researcher using the same personal computer during any given day. Instead, 2FA would be used to initially set up the communication between the researcher's computer and the PDC system. Thereafter, 2FA would only be needed to manage and revalidate the user's identity if the user wanted to change his or her password or connection information, or after too long a time had passed since the initial login, or if strange connection behaviour was detected.

For the sake of reliability and simplicity, the new PDC login solution should, wherever possible, rely on already existing software and not introduce new software tools. It is also important for the designed solution to be backwards compatible, so that experienced users who are comfortable with the existing Kerberos login method for PDC's soon-to-be-retired systems, Beskow and Tegner, can continue to log in in much the same way. Note that the existing Kerberos solution provides more protection for the current single-factor passwords than if those passwords were just being used over a remote SSH (Secure Shell) connection. However, the acquisition of Kerberos credentials will have to be migrated to 2FA once the new PDC login solution is phased in.

Based on all of these considerations, PDC is working on giving users a choice of two safe login options: an improved version of the login approach that has been used on PDC's resources for decades (without any major security incidents happening!), as well as a method based on registered SSH key pairs that will be in line with login mechanisms widely used on European systems. Both of these approaches will be made available to users via a new PDC web portal, where researchers will be able to initialise and update their login credentials for PDC's systems. Experience has shown that some PDC users regularly log in more than twenty times a day. For this reason, the new Dardel login approach (and the associated portal) is being designed so that the basic login process is as quick and hassle-free as possible, whether users log in to Dardel using Kerberos or SSH-key-based authentication. Further information about the improved login options and the PDC Portal will be available soon through the PDC Support pages. If you have any questions about the new login options, you can contact support@pdc.kth.se.

BioExcel News

Rossen Apostolov and Alessandra Villa, PDC

BioExcel Webinars

PDC hosts the popular BioExcel webinar series together with the University of Edinburgh. The webinar series (http://www. bioexcel.eu/webinars) features notable developments in the field of computational biomolecular research: success stories, new software releases, novel methods, new databases, and anything else that will be of interest and use to the research community. The webinars are held with a live audience so people can ask questions at the end of the presentations. The recordings of the webinars are then posted on the BioExcel YouTube channel (https://www.youtube.com/c/ BioExcelCoE) where the series has proven to be popular.

Last spring several of the webinars were by KTH researchers: Paul Bauer and Berk Hess gave a webinar on the new GROMACS release (https://bioexcel.eu/webinar-whats-new-in-gromacs-2021-2021-02-18), Berk Hess and Magnus Lundborg spoke about applying the AWH method to alchemical transformations (https://bioexcel.eu/webinar-applying-the-accelerated-weight-histogram-method-to-alchemical-transformations-2021-03-30) and Joe Jordan reported on the portable NonBonded-LIBrary NB-LIB (https://bioexcel.eu/webinar-nb-lib-a-performance-portable-library-for-computing-forces-and-energies-of-multi-particle-systems-2021-03-11). Join us for more interesting webinars in the autumn! The dates of upcoming webinars can be found on the BioExcel channels at https://bioexcel.eu.

Certificate of Excellence in RI Leadership

BioExcel has recently been awarded a Certificate of Excellence in Research Infrastructure Leadership for the participation of members of the staff team in the Executive Masters in Management of Research Infrastructures (EMMRI) at the University of Milano-Bicocca. The EMMRI programme is a general management qualification for science professionals with considerable work experience who are looking to further develop their managerial and strategic skills to lead research infrastructures. During the course work in the programme, the team developed specific use cases related to BioExcel Centre of Excellence operations and long-term sustainability and development plans.

PDC-Related Events

GROMACS Workshop

9-10 September 2021, online via Zoom

For details about this event run by BioExcel and the EuroHPC CC in Portugal, see https://www.eventbrite.pt/e/bilhetes-gromacs-workshop-162032753327.

HPC Sources

We recommend the following sources for other interesting HPC opportunities and events.

HPC in Sweden and Scandinavia
- SNIC
  https://snic.se
- SeRC
  https://e-science.se
- SeSE
  http://sese.nu
- NeIC
  http://neic.no
- ENCCS
  http://enccs.se

European HPC ecosystem
- HPC in Europe
  https://hpc-portal.eu
- EuroHPC
  https://eurohpc-ju.europa.eu
- PRACE
  https://www.prace-ri.eu
- LUMI
  https://www.lumi-supercomputer.eu
- ETP4HPC
  https://www.etp4hpc.eu
- EOSC
  https://eosc-portal.eu

A selection of projects that PDC is involved with
- BioExcel CoE
  https://bioexcel.eu
- TREX
  https://trex-coe.eu
- EBRAINS
  https://ebrains.eu
- PerMedCoE
  https://permedcoe.eu
- EXCELLERAT
  https://www.excellerat.eu
- EOSC-Nordic
  https://eosc-nordic.eu
- DICE
  https://www.dice-eosc.eu
- HPC-Europa3
  http://www.hpc-europa.eu

HPC news sources
- HPCwire
  http://www.hpcwire.com
- insideHPC
  https://insidehpc.com