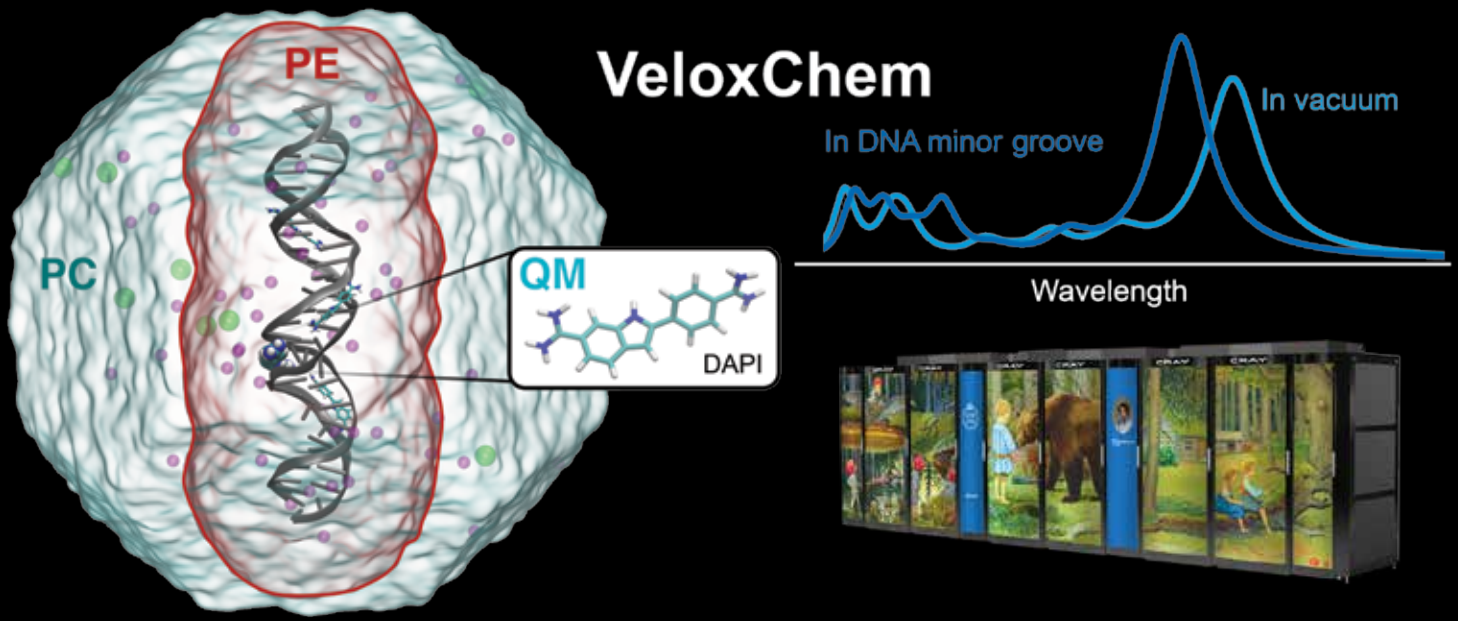
no.2 2019

PDC Center for High Performance Computing

PDC Newsletter



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: Erwin Laure
Editorial Staff: Genet Edmondson
Layout: Genet Edmondson
Graphic Design: Maria Malmqvist
E-mail: support@pdc.kth.se

ISSN 1401-9671

Cover

The cover image shows the calculated spectral shift of the DAPI biomarker as it binds strongly in the minor groove of adenine-thymine rich regions in DNA. The complex biochemical environment is modelled with the polarizable embedding approach in VeloxChem.

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

With the end of the year approaching, it is time to look back on our past achievements and to look forward to what the future will bring. For PDC, it will be an exciting future: SNIC has recently decided to fund a new national system at PDC that will replace Beskow in early 2021. Planning for the procurement of this system is already proceeding at full steam: we foresee a modular system, with a traditional, CPU-based, module to support most of the current workloads, as well as a forward-looking accelerated module paving the way to upcoming exascale machines, like in Europe, the US, and Japan. To provide Swedish users with the latest hardware, which is also being used in those systems, the second module will likely be added in 2022, while the first module should be available early in 2021 to facilitate a smooth transition from our current Beskow system.

The other major change PDC will see in 2020 is a change in leadership. After more than eleven years at PDC, it is time for me to move on. As of March next year, I will be joining the Max Planck Society as the leader of the Max Planck Computing and Data Facility (MPCDF) in Garching, Germany. This was not an easy decision as PDC and KTH have provided an amazing environment and I have had the opportunity to work with many excellent people – many thanks for this. I am sure we will keep in contact and hopefully also have future joint projects. While the search for a new PDC director is ongoing, Patrick Norman (who is the head of the Department of Theoretical Chemistry and Biology at KTH) will act as interim director to ensure the continued high quality of services provided by PDC.

To prepare for the future of supercomputing, particularly in the exascale era, massive investments are needed to make applications and codes fit for the changes to come. The cover article of this newsletter discusses the results of significant development work by the Department of Theoretical Chemistry and Biology at KTH. In collaboration with PDC, researchers from the department started a major effort to develop a new software platform for quantum molecular modelling, VeloxChem, which is designed for future HPC systems and the EuroHPC ecosystem. After two years of development, an initial version of VeloxChem, that already shows excellent performance and scaling behaviour, has recently been published.

Education and outreach continue to be major efforts at PDC and this autumn a number of high-profile events have been organized, including our yearly flagship PDC summer school, a PRACE/BioExcel school for biomolecular modelling, and an industry day to foster the uptake of HPC technologies in industry. We also contributed to

Below: Students enjoying the PRACE/BioExcel Seasonal School at PDC in June 2019

the further development of the European HPC training landscape through a workshop organized by the FocusCoE project. In addition, we started a new endeavour – the PDC Cafes: an opportunity to meet PDC staff in an informal setting, chat about recent developments in HPC, and solve problems you might have in using PDC services. Take advantage of this and stop by at some of the upcoming events!

With this, I would like to wish you all the best for the upcoming holiday season and let me express my deep gratitude and thankfulness for the excellent environment at PDC, KTH, and the overall Swedish research environment that I have had the honour to contribute to over the past eleven years. I look forward to seeing this wonderful environment develop further, particularly with the new national SNIC system at PDC. Let's keep in touch!

Erwin Laure, Director PDC

Did you miss the PDC-SeRC seminar on applying for large time allocations?

Early in October, Philipp Schlatter - who is the current chairperson of the Swedish National Allocation Committee (SNAC) - gave a seminar at PDC about the operation of the Swedish National Infrastructure for Computing (SNIC), which is responsible for providing large-scale computer resources for academic researchers in Sweden.

In particular, SNAC processes applications for large allocations of time for academic research. During the seminar, Philipp went through the specific procedures that are used for handling applications for large allocations. In case you missed the seminar, Philipp's slides are available on the event page (https://www.pdc.kth.se/about/events/pdc-serc-seminar-snic-and-snac-handling-of-large-scale-computer-time-allocations-in-sweden-1.924184).

In This Issue

Above: PDC-SeRC Seminar by Philipp Schlatter from SNAC, 4 October 2019

Staff Focus

Alessandra Villa

Alessandra Villa is an expert in solving biophysical problems using molecular modelling and simulations. After her doctorate in quantum chemistry at the University of Milan, Alessandra was awarded a Marie Curie Fellowship to work on biomolecular force fields and simulation at the University of Groningen in the Netherlands. After that she moved to Germany (where her work primarily focused on nucleic acid simulations) and then to Sweden where she worked at the Karolinska Institute

Alessandra has developed molecular models (both at the atomistic and coarse-grained levels) to better describe how biomolecules interact. She has worked on multiscale approaches to describe subcellular processes in neurons, with particular attention to the axon. Alessandra is a Docent in Biochemistry and she has taught at undergraduate and graduate levels. Alessandra joined PDC this September, as a part of the BioExcel European project. Her main task will be to work on user-driven-development focusing on the GROMACS software package.

When she has free time, Alessandra likes to read books and enjoy the arts or do outdoor activities. The seaside and mountains are her favourite landscapes.

VeloxChem — Quantum Molecular Modelling in HPC Environments

Zilvinas Rinkevicius, Xin Li, and Patrick Norman, Department of Theoretical Chemistry and Biology, School of Engineering Sciences in Chemistry, Biotechnology and Health, KTH

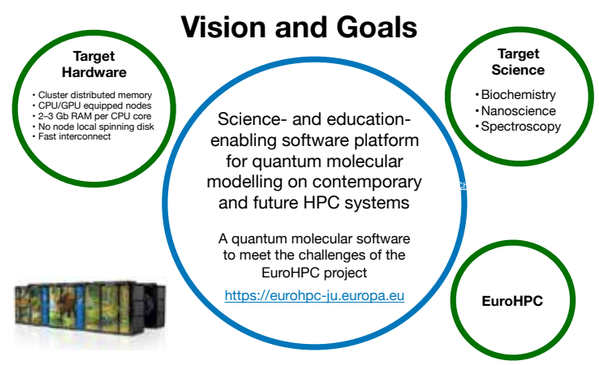
Vision and Goals

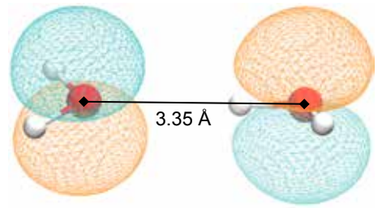
Theoretical chemistry has made tremendous progress in the past few decades and is today an indispensable tool in all fields of molecular science, exemplified by biochemistry and nanotechnology, where it can be employed to reveal the microscopic origins of functionality and interactions as well as to tune performance and guide synthesis. The interplay between simulation and experiment is a key enabling factor for advancements in the life and materials sciences, and spectroscopy represents one of the most important meeting points in between the two disciplines.[1] It is imperative to base spectroscopic modelling on first principles, as the underlying microscopic events are quantum-mechanical in origin[2] — although typically in combination with classical force field molecular dynamics for the sampling of configuration space. In order to study complex molecular systems and delocalised electronic transitions, it is consequently important that such first-principles approaches are developed into forms where they may be tractably applied to large systems. It is therefore not surprising that a vast amount of work in computational chemistry has been carried out to extend the computationally tractable system domain boundaries with respect to time and length scales.[3]

Alongside the development of methods and algorithms, we are witnessing an ever ongoing advancement of high-performance computing (HPC) cluster resources to reach higher numbers of floating point operations per second. In the first PDC Newsletter for 2019, PDC Director Erwin Laure presented the EuroHPC project and its leading Europe toward exascale computing (https://www.pdc.kth.se/publications/pdc-newsletter-articles/2019-no-1/eurohpc-europe-s-path-to-exascale-1.911735). EuroHPC comprises an entire supercomputing ecosystem with the ambition to strengthen European industry and academia alike, and as such it brings us an exciting future at the same time as it challenges us. A view of the most recent TOP500 list gives a clear indication that heterogeneous hardware solutions are to be expected when such extreme performance numbers are to be reached and there is a community consensus among quantum chemists that an efficient utilization of general-purpose graphical processing units (GP-GPUs) is difficult to achieve in our software where the compute-intensive kernels involve 10^4–10^5 lines of code. What we believe to be the competitive edge of the VeloxChem project in this respect is the fact that it started from scratch some two years ago without the burden of a code legacy.

Below: Vision and goal statement for the VeloxChem project together with an illustration of an educational aspect of Python-driven quantum chemistry software

```
from mpi4py import MPI
import veloxchem as vlx

# set up molecule, basis set and output stream
molecule = vlx.Molecule.read_str("""
   O       -2.55903     -0.54857     -2.45096
   H       -1.81050     -0.31857     -1.85689
   H       -2.24323     -0.20897     -3.30097
   O       -0.01061      0.30291     -0.45653
   H        0.94253      0.49302     -0.45202
   H       -0.22604      0.26272      0.49194""")

basis = vlx.MolecularBasis.read(molecule, "aug-cc-PVDZ")
ostream = vlx.OutputStream("water-dimer.out")

# run self-consistent field calculation
scf_drv = vlx.ScfRestrictedDriver(MPI.COMM_WORLD, ostream)
scf_drv.compute(molecule, basis)

# compute molecular orbital integrals in physicists' notation
moints_drv = vlx.MOIntegralsDriver(MPI.COMM_WORLD, ostream)
ijij = moints_drv.compute_in_mem(molecule, basis, scf_drv.mol_orbs, "OOOO")

# get Coulomb interaction between HOMO and HOMO-1
homo = molecule.number_of_alpha_electrons() - 1
print(ijij[homo, homo - 1, homo, homo - 1])

# generate cube files for HOMO and HOMO-1
cube_dict = {
   'cubes': 'mo(homo), mo(homo-1)',
   'files': 'homo.cube, homo-1.cube',
}
vis_drv = vlx.VisualizationDriver(MPI.COMM_WORLD)
vis_drv.gen_cubes(cube_dict, molecule, basis, scf_drv.mol_orbs, scf_drv.density)
```





The stated goals of the VeloxChem program (which are displayed in the figure above) encompass the following:

> *To deliver a science- and education-enabling software platform for quantum molecular modelling on contemporary and future HPC systems, and to be part of the EuroHPC ecosystem.*

Behind the term science-enabling there are a multitude of software requirements that we find important in our work, including (i) coverage of system sizes up to and beyond 500 atoms in the quantum region, (ii) accurate description of electronically excited states that show a more diffuse character than the ground state, (iii) stable and reliable convergence of iterative equation solvers also with use of diffuse basis sets, (iv) time-efficient prototyping of novel scientific approaches, (v) transparent exposure of data structures to enable in-depth analyses for standard users, (vi) flexible ways to interact with other components of the simulation (such as molecular dynamics, parameterizing the embedding, and data visualization) and, not least, (vii) a fast return of results so as to remain in synchronicity with experimental project partners. The term education-enabling adds another set of software requirements to this already long list. In this context, the notion of deeper learning refers to taking each student's understanding of the subject matter to another (deeper) level. Our experience tells us that the process of implementing methods to solve fundamental equations is supremely efficient as a means to achieve that type of deeper learning, but only a small number of students are granted this opportunity as many core modules of scientific software were written a long time ago and have often been made obscure by code optimization. What if we could instead offer access to the needed building blocks to explore quantum chemistry in very much the same manner that we can use the Python NumPy package to explore linear algebra?

All of these considerations went into the design of the VeloxChem program. We started out with the ambition to (i) scale standard spectroscopy calculations beyond 10,000 central processing unit (CPU) cores so as to be considered relevant

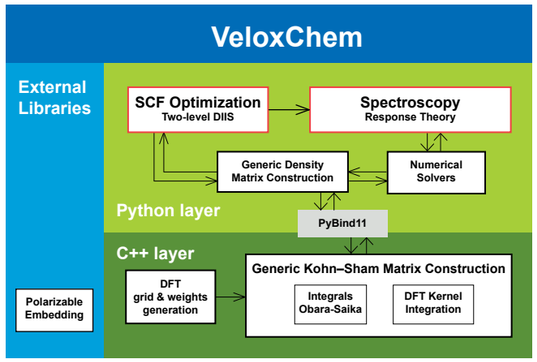
for the EuroHPC project and, at the same time, (ii) provide a platform for undergraduate students to implement (in Python, on a laptop) their own self-consistent field (SCF) optimizer of ground state wave functions within the limited time of a course exercise. We have now reached these goals and a more detailed presentation of the resulting VeloxChem program has been accepted for publication in the Wiley Interdisciplinary Reviews (WIREs).[4]

Program Structure and Modularity

From a programming language perspective, the VeloxChem program is designed in a two-layered fashion. The upper layer is written in Python and manages the use of the available hardware resources by means of dynamically split message passing interface (MPI) communicators for separate and parallel execution of the three compute-intensive components of the calculation namely (i) the analytical evaluation of electron-repulsion integrals (ERIs), (ii) the numerical grid-based integration of the Kohn–Sham density functional theory (DFT) exchange-correlation kernel, and (iii) the contribution due to the coupling between the quantum mechanical (QM) focus region of the system and the classical molecular mechanics (MM) embedding region (described by point charges and atomic-site polarizabilities). Any one of these three components will scale with the size of the hardware resource only up to a point and, by using split communicators and parallel execution, we can scale the calculation further and reach shorter wall times of execution. The splitting of the communicator is best made dynamic because every model system — depending on the relative sizes of the QM and MM regions, basis set, grid accuracy, and so forth — will experience different and hardware-dependent (CPU versus GPU) load balancing that needs to be adjusted for in order to avoid idling resources.

The Python layer also implements high-level quantum chemical methods such as (i) the direct inversion in the iterative subspace (DIIS) technique to reach the self-consistent

Below: VeloxChem program structure with (i) a Python layer implementing hardware resource management and high-level quantum chemistry and (ii) a C++ layer implementing the compute-intensive analytical integral evaluation and numerical DFT-kernel integration on a spatial grid



field (SCF) solution of the ground-state electron density and (ii) the several variants of iterative numerical response equation solvers that provide the spectral responses (see the figure above). Communication of data arrays between the Python and C++ layers is achieved with the header-only library Pybind11.[5] All underlying algorithms (SCF and response) are formulated such that there is one common set of data arrays that needs to be communicated through the layers, namely generalized density matrices from Python to C++ and the corresponding Fock matrices (after a contraction is made with the ERIs) from C++ to Python. However, this design also brings a more novel feature to the VeloxChem program as we may expose virtually any data array from the C++ layer to be accessed by the user as a NumPy[6] array in the Python layer. As a result, not only are students equipped with the tools to perform educational exploratory work as discussed in the introduction but general users can also perform in-depth analyses of their research. As an illustration of the latter aspect, the figure on the previous page shows how it is possible, with a Python code snippet, to set up the calculation of a water dimer system and obtain the Coulomb interaction energy between the electrons occupying the highest occupied molecular orbitals (HOMOs) of the respective water molecules. For the water dimer configuration under consideration, this interaction energy amounts to the following ERI:

$$J_{ij} = (ii|jj) = \frac{e^2}{4\pi\varepsilon_0}\int \frac{\psi_i^*(\mathbf{r}_1)\psi_i(\mathbf{r}_1)\psi_j^*(\mathbf{r}_2)\psi_j(\mathbf{r}_2)}{|\mathbf{r}_1-\mathbf{r}_2|}d^3\mathbf{r}_1 d^3\mathbf{r}_2 = 0.156543\ E_h$$

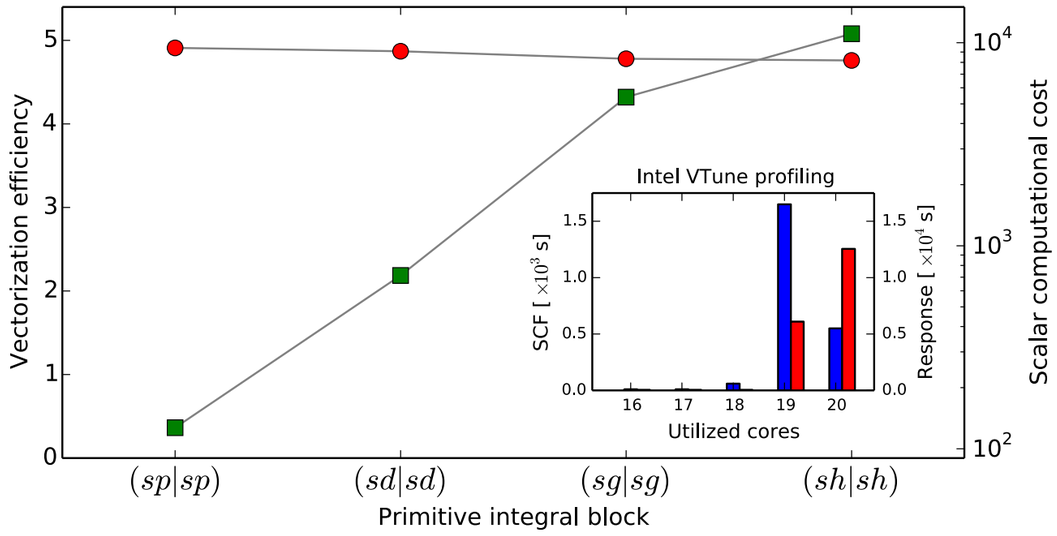
which, incidentally, is rather close to the interaction energy of 0.158 E_h for two elementary point charges separated by the O–O distance of 3.35 Å. We believe that every experienced computational scientist working with monolithic legacy codes can appreciate the ease with which the VeloxChem program provides transparent access to specific user-requested intermediate data in the calculation, extracted from arrays that, due to the sheer number of elements, would never be stored or made accessible to the user. At this point, the question to be asked is whether this code design comes at the price of reduced computational efficiency and we will address this question in the next sections.

Vectorization and Parallelization

VeloxChem employs a modified version of the Obara–Saika scheme[7] to analytically evaluate ERIs. The implementation is based on automated code generation with a high degree of loop structure optimization in order to achieve efficient vectorization on CPUs that implement single instruction multiple data (SIMD) instructions. The obvious goal of this vectorization is to achieve efficient usage of the vector registers on modern CPUs. In the figure below, we present a performance profile of the vectorization for the vertical recursion step in the workflow scheme. For basis set primitives up to angular momentum l = 5 (or h-functions), the efficiencies reported by the Intel-compiler and depicted in the figure exceed a speedup of 4.75 for the vectorized code as compared to the scalar code.

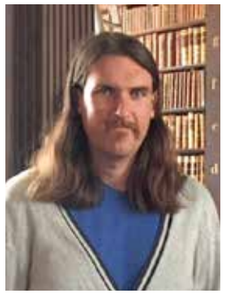
Vectorization Efficiency

The OpenMP parallelization on a single node is benchmarked using the Intel VTune Amplifier tool that measures the amount of time spent at different levels of CPU core activity. Thus, the range of activity spans from zero (fully idling CPUs) to the maximum number of cores on the node (in our case 20). The inset in the figure below depicts the reported results for the complete SCF optimization of



Above: This figure shows SIMD vectorization for primitive integral blocks with different angular momenta (red circles) and the corresponding scalar computational costs (green squares) to be understood in a relative sense. The performance profiling results in the inset are obtained for the C_{60} fullerene at the level of Hartree–Fock/def2-SVP and recorded with the performance profiler Intel VTune Amplifier running on a dual-socket (20-core) Intel Xeon node that implements AVX2.

Staff Focus



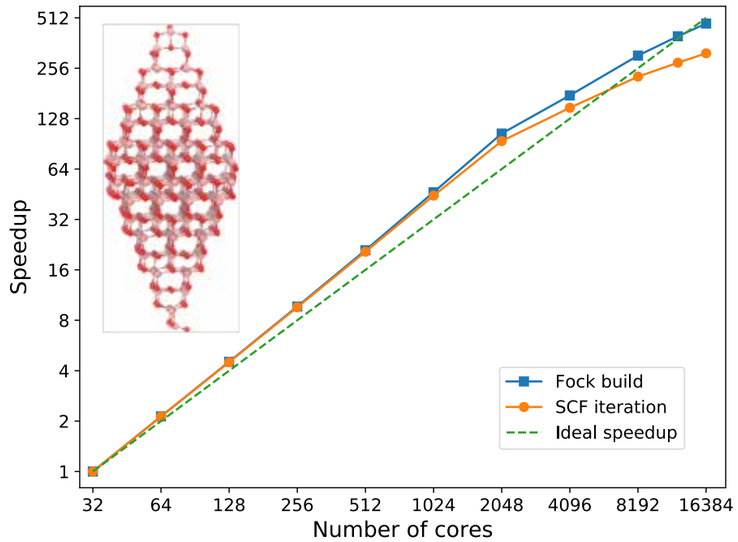
Joe Jordan

Joe got a bit of a late start as an academic, entering university at age 24 and finishing three years later with two bachelor's degree, one in biochemistry and another in computational mathematics from The University of Texas at Austin. During this time his interests in research were sparked by his work in organic and bio-organic synthetic chemistry. Joe then went on to pursue doctoral studies at the University of Pennsylvania in biochemistry and molecular biophysics. His projects revolved around understanding the effects of cancer mutations to better facilitate personalized cancer treatment. The methods he applied spanned from molecular dynamics (MD) simulations of proteins to machine learning and database mining. He also worked on a project which resulted in a graphical user interface for the GROMACS MD simulation package. After completing his Ph.D., Joe went on to do post-doctoral research at Stockholm University, where he was involved in a number of projects that aimed to make it possible to use experimental data to guide MD simulations.

Currently, Joe is working as a research software engineer at PDC. His project

…continued on page 13

Below: This graph shows scaling with respect to the number of cores for the SCF optimization of the ground state wave function. Speedups for one complete SCF cycle (orange circle) and the isolated associated Fock matrix construction (blue square) are reported separately and with reference to the calculation using 32 cores. Results are obtained for a titanium oxide nanoparticle, $Ti_{165}O_{330}$, at the level of Hartree–Fock/def2-SV(P), resulting in 8,580 contracted and 18,810 primitive basis functions.



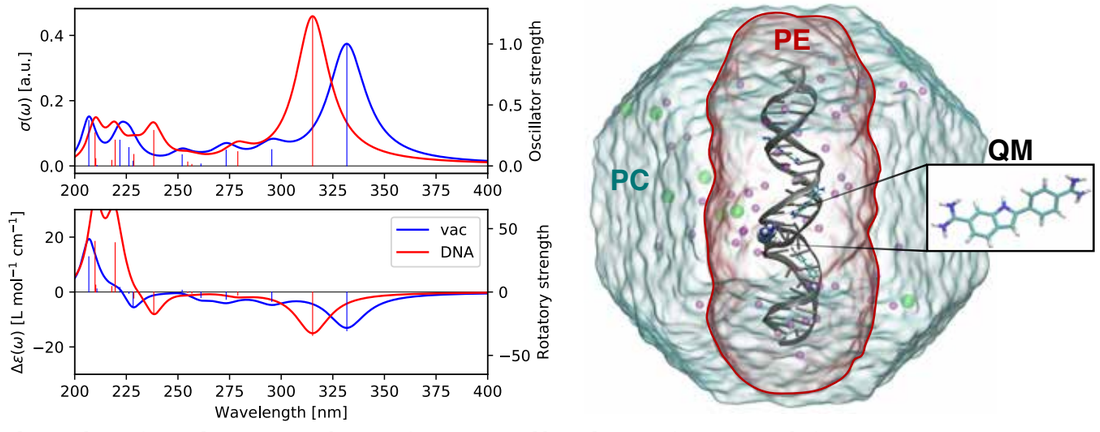
the Hartree–Fock ground state of the fullerene C_{60}, as well as the subsequent complete response calculation to obtain the five lowest roots of the generalized eigenvalue equation, that is, the five lowest singlet states in the UV/vis spectrum. The overall total OpenMP threading efficiency is seen to reach levels of some 96% and 98% for the SCF (blue bars) and response parts (red bars), respectively.

In the next figure, we address a titanium oxide nanoparticle that illustrates the size (but not necessarily type) of system that we routinely wish to encompass in the quantum region of a given physical model. With generous and exclusive access to a large part of the main cluster at the PDC Center for High Performance Computing in Stockholm, we had the opportunity to test the MPI/OpenMP scaling of VeloxChem up to 16,384 cores (or 512 dual socket 32-core Intel Xeon nodes) and we used the otherwise identical calculation performed on 32 cores (or 1 node) as the benchmark reference. A complete SCF cycle during the optimization of the ground state density requires a wall time of 70,565 s in the reference calculation, out of which 70,493 s and 72 s are spent on the construction of the Fock matrix and other parts (including the Fock matrix diagonalization) respectively. Apart from the Fock matrix construction, the SCF module is written in Python (as shown in the figure on page 6) and executes on a single node and the corresponding timing is therefore nearly constant along the series of calculations presented in the above figure.

Owing to favourable differences in the dynamic workload balancing between the available MPI processes and OpenMP threads, the Fock matrix construction scales slightly beyond linear with respect to the number of cores (up to a point of 384 nodes) and reaches an efficiency of 162.9% at the point of using 64 nodes and compared to the reference calculation. This example thus demonstrates almost perfect scaling up to the point of allocating almost one CPU core per primitive basis function. The wall time for the Fock matrix construction using 512 nodes is reduced to 149 s, so while it may be possible to observe a reasonable scaling beyond 512 nodes in the present example; it makes little sense in practice as the SCF/DIIS part becomes comparable in duration.

Spectrum Calculations

As a single example of the scientific capabilities of VeloxChem, the figure below shows the calculated UV/vis and electronic circular dichroism (ECD) spectra of the DAPI biomarker inside the minor groove of DNA as studied in an earlier paper.[8] The model system illustrated in the figure is set up with (i) a DAPI biomarker in the QM region, (ii) a polarizable embedding (PE) region containing the DNA sequence, the three additional bound DAPIs, and a 15 Å shell of water molecules with a small number of sodium and chlorine ions (the entire system is charge neutral), and (iii) the remaining water molecules and ions in the surrounding point-charge (PC) region. In total, the model includes some 85,000 and 23,000 point-charge and polarizable sites, respectively.



Above: Absorption and ECD spectra for DAPI in vacuum and bound to DNA in aqueous solution
The results were obtained at the level of TDA/BHandHLYP/6-31+G(p,d). Oscillator strengths and rotatory strengths (in units of 10^{-40} esu²cm²) are broadened by a Lorentzian line profile with an HWHM of 0.124 eV to produce the displayed spectra.

Outlook

With the publication of the program release paper[4], VeloxChem version 1.0 will be distributed under the GNU Lesser General Public License version 2.1 (LGPLv2.1) licence and made available for download from the VeloxChem home page https://veloxchem.org. At present, VeloxChem offers spectrum simulations of complex molecular systems on CPU-based HPC cluster facilities. With a layered Python/C++ language design, it achieves excellent MPI/OpenMP parallel scaling and SIMD vectorization for realistic systems and using up to a quarter of the Beskow cluster. The compute-intensive integral code is large but remains agile and adaptable to new hardware solutions owing to the fact that it is based on automated code generation.

A main future focus point is to develop CUDA implementations of the three compute-intensive components of VeloxChem namely (i) analytic evaluation of integrals, (ii) numerical DFT kernel integration, and (iii) polarizable embedding. Out of these three, the implementation of integrals is design-wise the difficult one. We are grateful for the kind invitation to participate at the Eurohack19 workshop in Lugano (which ran from the 30th of

September to the 4th of October 2019), where we received general guidance and dedicated expert help to start on this endeavour. During this intense week, we implemented a first version of a code for integral evaluation in the primitive Gaussian orbital basis with Cauchy–Schwarz screening, and we saw tentative performance numbers that point toward a great future for VeloxChem in the landscape of heterogeneous cluster computing.

Acknowledgements

The VeloxChem project is enabled by our participation in the Swedish e-Science Research Centre (SeRC) and by integrated collaboration with the PDC Centre for High Performance Computing. The results presented in this article were obtained with the use of computational resources provided by the Swedish National Infrastructure for Computing (SNIC).

References

1. Lindon, J. C., Tranter, G. E., Holmes, J. L., Eds. *Encyclopedia of spectroscopy and spectrometry*; Academic Press: San Diego, 2010.
2. Norman, P.; Ruud, K.; Saue, T. *Principles and practices of molecular properties*; John Wiley & Sons, Ltd. Chichester, UK, 2018.
3. Akimov, A. V.; Prezhdo, O. V. *Large-Scale Computations in Chemistry: A Bird's Eye View of a Vibrant Field*. Chem. Rev. 2015, 115, 5797–5890.
4. 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. 2019, DOI: 10.1002/wcms.1457.
5. Jakob, W.; Rhinelander, J.; Moldovan, D. *pybind11 – Seamless operability between C++11 and Python*. 2017; https://github.com/pybind/pybind11.
6. van der Walt, S.; Chris Colbert, S.; Varoquaux, G. *The NumPy array: A structure for efficient numerical computation*. Comput. Sci. Eng. 2011, 13, 22–30.
7. Obara, S.; Saika, A. *Efficient recursive computation of molecular integrals over Cartesian Gaussian functions*. J. Chem. Phys. 1986, 84, 3963–3974.
8. List, N. H.; Knoops, J.; Rubio-Magnieto, J.; Idé, J.; Beljonne, D.; Norman, P.; Surin, M.; Linares, M. *Origin of DNA-Induced Circular Dichroism in a Minor-Groove Binder*. J. Am. Chem. Soc. 2017, 139, 14947–14953.

Nobel Calling at PDC

Thor Wikfeldt, PDC

Since last year, the Nobel Prize Museum has run an event, known as Nobel Calling Stockholm (https://nobelprizemuseum.se/en/nobel-calling), during "Nobel Week", that is, the week when all the Nobel Prize laureates are announced. Nobel Calling is a series of events involving organisations and institutions in Stockholm – it focuses on the importance of science, literature and efforts to bring peace to the world. As part of this event, the KTH Royal Institute of Technology (KTH) offers tours of research labs so people can see research environments in areas such as self-driving vehicles, robotics and electricity.

This year PDC was again asked to host a group of inquisitive people who wanted to learn about some of the research infrastructure available at KTH. Around 20 people from different walks of life turned up. They were given an introduction to what supercomputers are and what they are used for and then taken on a guided tour of the PDC machine room. After reconvening outside the machine room (where it is less noisy), questions abounded on all kinds of topics related to PDC and supercomputing in general!

Above: PDC was one of the labs people could visit during the KTH Lab Tours on 9 October 2019 as part of the Nobel Calling Stockholm event.

Below: Preparing for the fire test by setting up a plastic door into a computer hall with a fan to increase the air pressure in that hall

Testing Our Gas-Based Fire Suppression System

Gert Svensson, PDC

In October 2019 we performed a test of the air-tightness of PDC's three computer halls (the main hall, the tape-robot room and the small computer hall). The background to these tests is that we use a gas-based fire suppression system in these rooms. If a fire is detected in one of our computer halls, a mixture of the inert gases argon and nitrogen (sometimes referred to as Argonite) is released into the room from a bank of high-pressure gas bottles (shown below). The idea is to reduce the oxygen level in the room from the normal 18% to around 11%. This is enough to extinguish the fire, but humans can still breathe, at least for a short time. (The sensation would be comparable to breathing when very high up on a mountain.) Of course, a warning is sounded before the gas is released so people can exit the room. When the gas comes into the room at a high pressure, the pressure in the room would rise enough to destroy doors, windows and even walls if special measures were not taken. In our case, special pressure equalization dampers are opened automatically. Before the gas is released, fire dampers shut off the normal ventilation to the room. After the gas has entered the room, the pressure equalization dampers are closed again to keep the gas in the room for as long as possible. It is this step where the air-tightness of the room is highly important. If the gas is diluted too much, the fire could re-ignite. The fire code states that the gas should be kept at an effective level for at least ten minutes.

In the test, the normal door to the room is replaced by a frame with air-tight plastic and with a big fan that can increase the pressure in the room (see images above) and sensors that can measure how fast the pressure drops. In our case, we found that the rooms were more air-tight than required. However, we also found that, for the small computer hall, too much gas was released for the size of the room. This must have been caused by a human error in the design of the system in 2003. Whether this is within the safety limits is not yet known at the time of writing.

Above: (left) Rack of high-pressure gas bottles containing Argonite to suppress fire breakouts (centre) The sampling smoke detector sucks in air from the red pipe which has small holes at regular intervals and uses a laser detector to measure the smoke level in the room. (right) The nozzles where the gas enters the room are equipped with silencers to reduce the sound level when the gas is released. This is important as modern hard disks are sensitive to (and could be damaged by) high sound levels.

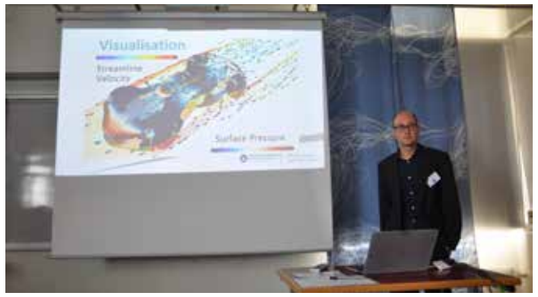
PDC Industry Day 2019

Lilit Axner, PDC

On the 19th of September, PDC held its second PDC Industry Day, which has now become an annual event for PDC. This year's event was a great success with over 30 participants from 12 companies that represented large industries, small and medium-sized enterprises (SMEs) and independent software vendors (ISVs). These were Scania, Sweco, Tyréns, ABB, Briggs Automotive Company (UK), Airinnova, EDRMedeso, ESI, Icarus Digital Math, Volupe, Berotec and Neptech.

The PDC director, Erwin Laure, opened the day and discussed current and future developments in high-performance computing (HPC) towards exascale. The presentations at the event were about the use of HPC in engineering, covering a broad industrial research spectrum including aircraft design, speed of embankments for high-speed railway, simulation of race supercars, large track design, hydrodynamic simulations in dam safety assessments and more.

The invited keynote speakers were Greg Cartland-Glover from UKRI STFC - Briggs



Above: Keynote presentations by Gregory Cartland-Glover from UKRI STFC - Briggs Automotive Company, UK (top) & Mattias Chevalier, Scania, Sweden (bottom)

Below: Pierre-Louis Ligier from Sweco Energy, Sweden

Automotive Company and Mattias Chevalier from Scania. Mattias discussed using Beskow to simulate heat transfer in Scania trucks (rather than performing a large number of physical experiments) and thereby saving substantially on development costs. Greg discussed a useful PRACE SHAPE project UKRI STFC undertook together with Briggs Automotive on the Scottish supercomputer Archer at EPCC. The third PDC Industry Day is planned for September 2020. To attend it, and similar PDC events, please check: https://www.pdc.kth.se/about/events for details.

NORDIC E-INFRASTRUCTURE COLLABORATION

NeIC Newsflash: Annual Open Call & 2020-2025 Strategy

Michaela Barth, PDC

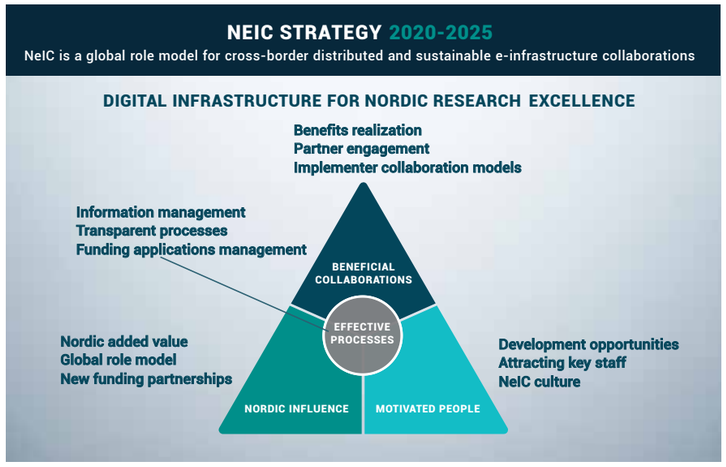
NeIC, the Nordic e-Infrastructure Collaboration, was established in 2012 under NordForsk with the joint operation of a high quality and sustainable Nordic Tier-1 service within the Worldwide Large Hadron Collider Computing Grid (WLCG) as the single core activity. Since then, NeIC has grown to encompass a broad range of scientific application areas as areas of collaboration on e-infrastructures. This came about through the implementation of the Nordic eScience Action Plans and from responses to Open Calls for Letters of Interest by national e-infrastructure providers and Nordic research groups with clear e-infrastructure needs.

In May 2019, NeIC issued an Open Call (https://neic.no/news/2019/05/23/open-call): consortia consisting of e-infrastructure providers, developers, researchers and related research communities around the Nordic region were invited to propose collaborations of joint Nordic interest within the area of digital infrastructure. Five project proposals and two pre-study proposals were received. The decision on the pre-studies was made in early September and one pre-study on a Nordic Digital Humanities Laboratory (https://neic.no/ndhl) has started. The project proposals are being reviewed by an independent expert panel. The final decision, based on the resulting suggested ranking, will be made by the NeIC Board at their mid-December meeting, which is likely to lead to one successful project being started in 2020.

One of the criteria for a successful project application is compliance with the NeIC strategy. The vision for NeIC is to be a global role model for cross-border distributed and sustainable e-infrastructure collaborations. NeIC's renewed strategy for 2020-2025 (see figure below) emphasizes beneficial collaborations, motivated people, efficient processes and Nordic influence. According to this strategy, NeIC activities will centre on digital infrastructure for Nordic research excellence.

The intention is to establish the open call as NeIC's main tool for project creation in the future. A predictable, fine-tuned annual open call process is expected to even out the budget situation over the coming years so that there are not too many projects starting and ending all at the same time. Next year's round is intended to open earlier to make it possible for successful projects to be notified in autumn 2020.

Feedback on the Open Call's scope and improvements for next year's round of calls are welcome: contact us at neic-helpdesk@neic.no with your suggestions!



Above: NeIC strategy for 2020-2025

Staff Focus

…continued from page 8

is to expose the most computationally intensive aspects of MD simulations, namely the non-bonded calculations, as a library that any simulation engine can link against. This will enable external codes and researchers to get the great performance that GROMACS is known for without needing to understand how GROMACS handles data or hardware. This should help researchers focus more on their domain of expertise while also making it possible for them to make more efficient use of compute resources.

Above: On the 8th of October 2019 an "Introduction to PDC Systems" course was held at PDC. If you would like to learn how to use our systems for your research, please keep watching our Event calendar (https://www.pdc.kth.se/about/events) and attend the next time the course is run (in spring 2020).

Staff Focus

Karl Johan Westrin

Karl Johan Westrin is a Ph.D. student in bioinformatics at KTH. Most Swedish doctoral students are required to spend 20% of their Ph.D. time gaining practical work experience (known as "*institutionstjänstgöring*" in Swedish), usually by assisting with university teaching. However, unlike many Ph.D. students, Karl Johan is working at PDC in first-line support rather than teaching. Karl Johan says that being able to undertake his Ph.D. work duties as a PDC support technician gives him "the best of two worlds".

Karl Johan has an M.Sc. in Computer Science from KTH, and worked as a back-end developer and release manager at a couple of private companies, most prominently the e-advertising company Cxense and the financial technology startup Aphelion, before embarking on his research journey through studying cone setting in the European spruce (*Picea Abies*) by reconstructing and analyzing related transcriptomes.

In his few precious minutes of spare time, Karl Johan works on a project for tune-to-lyrics-matching (and vice versa). He also likes to read (and to write!) Esperanto, and to play floorball at KTH-hallen (almost) each Friday.

Below: PDC Summer School, 19-30 August 2019

PDC Summer School 2019

Stefano Markidis, CST

There were 62 participants at this year's PDC Summer School. They were a mix of KTH M.Sc. students, Swedish and international Ph.D. students, and postdoctoral researchers. The international attendees came from institutions in Germany, Italy, Norway, Spain and Switzerland.

In addition to the usual lectures on MPI, OpenMP and CUDA, this year's summer school included lectures by Prof. David Broman introducing computer architectures, and lectures on performance monitoring, GitHub and containers by PDC staff members.

Above: PRACE/BioExcel Seasonal School, 10-13 June 2019

PRACE/BioExcel School

Lilit Axner, PDC

This year, for the second time, PRACE and the BioExcel Centre of Excellence organized the HPC for Life Sciences Seasonal School. The school was held from 10-13 June 2019 at the KTH Royal Institute of Technology in Stockholm and was another great success with 68 participants from 15 countries: Sweden, Norway, Denmark, Germany, the Netherlands, Portugal, Italy, France, Croatia, Hungary, India, the Russian Federation, Turkey, the UK and the USA.

This year's summer school covered the most popular codes in Molecular Dynamics, namely GROMACS, NAMD, VMD and AMBER, which were presented by the developers themselves. However, in contrast to the school held two years ago, this seasonal school was intended for advanced users and hence detailed more in-depth aspects of usage and development relevant to these codes. During the hands-on sessions in the afternoons, participants were able to test the advanced-level methods that were presented during the morning sessions. As usual, the PDC Support team and Application Experts were present during the whole school to assist the participants to use the Beskow Tier-1 system at PDC. Beskow is the Tier-1 system from Sweden that is integrated into the PRACE infrastructure.

During the four days of the school, there were two social events: a get-together dinner at the Syster O Bror restaurant on the KTH campus, as well as a nice boat trip through the mesmerizing Stockholm archipelago, which included a typical Swedish dinner buffet. The PRACE/BioExcel School was again an unforgettable experience that combined international networking, high-level expertise and knowledge transfer, as well as pleasant moments of social events. To participate in similar schools and training in the future, you are welcome to follow the PDC training calendar at https://www.pdc.kth.se/about/events/training, the PRACE calendar of training at http://www.training.prace-ri.eu and the BioExcel one at http://www.bioexcel.eu.

Below: Solving problems during labs at the PRACE/BioExcel Seasonal School made everyone hungry!

PDC Cafes

Thor Wikfeldt, PDC

Everyone working with technical user support knows that some issues are just so much easier to talk about in person rather than over email. Moreover, high performance computing (HPC) users may also be hesitant to ask certain general questions over impersonal email. With this in mind, and to further improve PDC's user support, we came up with a new type of recurring event this autumn: the PDC Cafe. The idea of the PDC Cafe is to create an informal meeting point between current and prospective HPC users and PDC support staff, where any issues related to HPC in general, and PDC services in particular, can be discussed. Topics can range from improving batch scripts or parallel performance to general questions on software management or how PDC's services can be improved to meet user needs. We are holding cafes roughly once a month: the first PDC Cafe was held on the 10th of September, and the second on the 16th of October. So look out for announcements for upcoming PDC Cafes in the new year: you are welcome to join us for a friendly chat about improving your research over a cup of coffee or tea and biscuits!

FocusCoE: Very Successful European HPC Training Stakeholder Workshop

Lilit Axner, PDC, with special thanks to Simon Wong, ICHEC

On the 8th of October 2019, the Horizon 2020 project FocusCoE, in collaboration with the European Commission, ran a European HPC Training Stakeholder Workshop. The purpose was to gather experts and stakeholders to define the high performance computing (HPC) training requirements for different domains and target audiences across academia and industry, and to examine how those requirements may be met by existing or new European and national programmes on HPC education and training. The workshop attracted 32 participants from institutions in 14 European countries, who represented the interests of different stakeholders including HPC users and developers, along with both academic and business education/training providers. Over three thematic sessions (which focused on different research communities and career progression), the participants and panellists discussed various training methodologies (like face-to-face and online learning), initiatives (such as mobility and certification programmes) and best practices (for example, engagement with small to medium-sized enterprises (SMEs) and integration into university programmes) to enhance awareness of, and the productive adoption of, HPC technologies in the advent of European exascale capability being realised via the EuroHPC Joint Undertaking.

The FocusCoE project aims to create a platform, the EU HPC CoE General Assembly, that enables all HPC Centres of Excellence (CoEs) to collectively define an overriding strategy and collaborative implementation for interactions with and contributions to the EU HPC Ecosystem. It supports the HPC CoEs to achieve enhanced interaction with industry, and SMEs in particular, through concerted outreach and business development actions. FocusCoE also instigates concerted action on training by and for the complete set of HPC CoEs: providing a consolidating vehicle for the user training offered by the CoEs and PRACE, as well as providing relevant training for the CoEs (for example, on sustainable business development).

FocusCoE promotes and coordinates the capabilities of, and services offered by, the HPC CoEs and the development of the EU HPC CoE brand, thus raising awareness with stakeholders and both academic and industrial users.

PDC-Related Events

PDC Summer School 2020

17-28 August 2020, KTH, Stockholm

For details of the next summer school, watch https://www.pdc.kth.se/about/events.

HPC Sources

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

BioExcel
https://bioexcel.eu/news-and-events/events

CERN
https://home.cern/scientists/events/computing

EGI
https://www.egi.eu/category/events

HPC University
http://www.hpcuniversity.org/events/current

HPCwire
http://www.hpcwire.com/events

NeIC
http://neic.nordforsk.org

PRACE
http://www.prace-ri.eu/HPC-access
http://www.training.prace-ri.eu
http://www.prace-ri.eu/events
http://www.prace-ri.eu/news

SeSE
http://sese.nu

SNIC
http://www.snic.se/news-events
http://docs.snic.se/wiki/Training

XSEDE
https://www.xsede.org

Nobel Calling at PDC

Learning about the PDC research environment during a KTH Lab Tour on 9 October 2019 as part of the Nobel Calling Stockholm event