NSC News
National Supercomputer Centre at Linköpings universitet
2013:1

# Triolith to be expanded


## Time is of the essence

Petros Souvatzis at Uppsala University use massively parallel calculations on Triolith to study material science.

Read further on page 3.

## Remote desktop visualization

ThinLinc provides improved performance for graphical applications on Triolith.

Read further on page 5.

## Triolith to be expanded

The NSC system Triolith inaugurated on January 24 will be expanded with an additional 320 nodes during the summer.

Read further on page 6.

# Science Enabling

*The hallmark of NSC is, and shall always be,* ***science enabling****. While this is an ambition and goal that is easy to set, it requires dedication in issues from small to large to fulfill and it brings a need for quality assurance (QA) to be convincing. To some extent QA is already in operation in our day-to-day activities by,* ***e.g.****, monitoring of support ticket alertness and cluster usage and uptime, but* ***science enabling*** *means so much more. We try to take an overviewing perspective on the production environment that you, the users, meet when using our services. This consideration includes so many aspects but one of my personal benchmarks bottles down to the time it takes from the point of arrival of an external scientist in your local research group until he/she compiles programs and launches calculations, billing core hours on your NSC project. NSC has streamlined this process to an extent that truly* ***enables*** *short-term scientific missions for researchers. The production environment is under constant development and improvement. One of the new things that we are particularly proud to announce is our remote ThinLinc visualization service that allows you to run a large selection of visualization software on our graphics accelerated servers and display the resulting images on your local screen over a modest-speed connection. This relieves the need for transfer of large data files before visualizing your simulation results. And let me also mention that we do our utmost when it comes to security surveillance and keeping your data and compute environment safe from intruders and misuse. These are all tangibles in support for our claim of being* ***science enabling****. But this editorial column intends also to bring up a somewhat more controversial argument namely our flagship computer Triolith itself, which recently has been inaugurated and presented as Sweden's most powerful computer earning position 83 on the Top500 list. But, in my opinion, this is a severely limited measure of such a national investment. Would NSC have done better if we, with the same amount of money, had landed a computer at position 75, or even higher? There is no telling, because we strive for science enabling and not peak flop capacity per se. In comparison, a scientist is accustomed to a constant and diverse evaluation. There is no end to the indices designed to enable fair and square comparison of scientists within and across research fields before handing out research means. So how can it be that we do not have index evaluations made of our large national investments in e-infrastructure. It may not need to be much more complicated than an index I given by*

$$I = \frac{G}{x+y}\sum_{i=1}^{N} f(i),$$

*where the impact factors f are summed for all N publications to which the computer has contributed. The sum is multiplied by* **G**, *the number of research groups behind the publications, and divided by x+y, the sum of investment and running costs of the computer. This is the kind of assessment of Triolith that I would welcome, and, ultimately, it would become a judgment of NSC's ability to make the most of the money. I believe Triolith would stand tall in any such comparison and thus support NSC's claim to be* ***science enabling****.*

PATRICK NORMAN, NSC DIRECTOR

## Kappa to be retired

It is time to retire the much appreciated and used NSC workhorse Kappa. Kappa arrived at NSC in October 2009 and was taken into full regular production in April 2010. During it's three years of operation 170 projects have had time allocated on the system with 600 unique users this far. The availability have been about 99 % and the average utilization about 90 %. A somewhat low figure compared to other NSC systems but explained by the special queue configuration with guaranteed nodes and risk jobs on part of the Linköping University specific par of Kappa.

The SNAC part of Kappa will be retired on July 1, 2013 and the part exclusive for users from Linköping University is planed to be retired on January 1, 2014.

Allocations on the SNAC part, that is SNAC Large, Medium and Small Scale projects will have their allocations transferred to the NSC resource Triolith. They will receive an allocation on Triolith of 25 % of their Kappa number of core hours from June 1, 2013 to be able to start migrating, and of 75 % of the Kappa number of core hours from July 1, 2013 until the end of the project. One Triolith core is about twice as fast as one Kappa core and 75 % of the Kappa allocation is thus an increase of resources.

Projects with allocations on the Linköping University part of Kappa will have to apply for other resources, such as Triolith, in order to replace their Kappa allocations from January 1, 2014

PETER MÜNGER, NSC

# Time is of the essence

Ever since computational materials set sail in the early sixties, the predictive power and efficiency of materials calculations have developed in close connection with the hardware evolution. The necessity of computational power could easily be understood by having a closer look at the minimum number of atoms needed for a quantitative study of, for example, the thermal expansion of a metal. In order to succeed with this theoretical endeavor, the movement of around 100 atoms needs to be simulated, and for these simulations to be finished in a foreseeable time, one needs to perform these calculations in parallel on equally many computational cores.

Another example of the need of computational power is the first principles prediction of the dependence of the vibrational lifetimes in a crystal upon the vibrational wavelength. Here the minimum number of atoms needed in a simulation is around 500–1000, requiring once again an equal amount of computational cores to be used in parallel. Here it is worth mentioning the calculations of vibrational lifetimes in the metals Li, Na, Ti and Zr, purely from quantum mechanical principles, made possible by the use of the NSC computational clusters (Figure 1). (J. Phys.:Condens. Matter. 445401, **23** (2011) using the so-called Self Consistent Ab Initio Lattice Dynamical (SCAILD) method. (Phys. Rev Lett. **100**, 095901 (2008)).

But not only the study of atomic degrees of freedom can benefit from massive parallel computing, it is also essential to gaining insight about how the electronic degrees of freedom effect materials properties. Here as second example we present a very accurate quantum diffusion Monte Carlo calculation (C. J Umrigar et al, J. Chem,. Phys. **99**, 2865 (1993)) of the charge density of water (Figure 2) done by the Uquantchem code on 960 of the computational Cores on the Triolith cluster.

The Diffusion Quantum Monte Carlo (DQMC) algorithm is built upon the so called path integral formulation of quantum mechanics, where as opposed to classical mechanics where particles only "travel" along singular paths in space and time, all possible paths the particles can undertake have to be taken into account in order to calculate the ground state energy and the corresponding charge density. Since these possible paths are independent of each other, the computation of their corresponding contribution to the total energy and charge density of the system can easily be parallelized by calculating one path per computational node. This is the basis of the parallel implementation of DQMC in the uquantchem code and the reason why such an implementation is so well suited to run on a massively parallel cluster such as Triolith.

The final examples are from some Benchmark calculations utilizing Second order Möller-Plesset perturbation theory (MP2) and Configuration Interaction Singles and Doubles (CISD), as implemented in the uquantchem code. These computational approaches are similar to DQMC in that they try to take into account the effects of electron-electron correlation, however since the theory of these methods differ fundamentally from that of the DQMC method, so does also the implementation of the parallelization. For instance in the case of the MP2 parallelization, it has been performed over the sum of double excitations used to calculate the energy correction to the electron correlation energy, whereas in the CISD case, both the construction of the Hamiltonian matrix (matrix describing the physics of the system) and the diagonalization of the matrix have been parallelized. Hence, the much more drastic speedup as a function of processors used , compared to the speedup obtained in the MP2 type of calculations (see Figure 3). Furthermore, thanks to the "divide and conquer" SCALAPAC (Scalable Linear Algebra Package) diagonalization routine PDSYSEV, the storage of the Hamiltonian matrix is shared over the memory of all participating nodes, making it possible to avoid major memory bottlenecks.

## New staff member

I moved at NSC in January as an application expert in climate modelling data, after previously working for four years in Stockholm within the Bert Bolin Centre for Climate Research. Before that I worked for some years in my home country of New Zealand as a climate modeller for the National Institute for Water and Atmospheric Research (New Zealand's equivalent of SMHI). I am very happy to be part of the NSC team and look forward to helping climate modellers from around Sweden to make the best use of NSC resources for their research.

HAMISH STRUTHERS, NSC

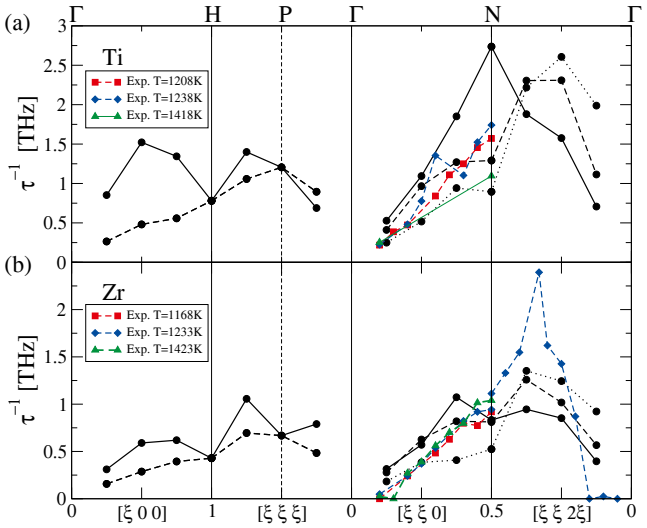



Figure 1: The computed atomic vibrational lifetimes in the Body Centered Cubic (BCC) crystal of Titanium and Zirconium, along different directions in the crystal. Here non-zero numbers at the horizontal axis correspond to finite wavelengths of the vibrations, whereas zero correspond to infinite wavelength.

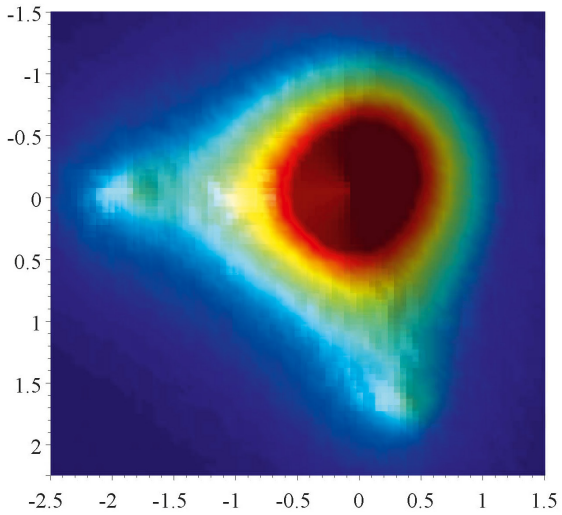


Figure 2: The computed charge density in water using a diffusion quantum Monte Carlo algorithm as implemented in the Uquantchem code. Here red colors correspond to high values of the charge density and blue or bluish colors correspond to low values. The units on the horizontal and vertical axes are atomic units.

Thus as seen from the above examples, when time is of the essence in science, the implementation of parallel computational software in close connection to powerful computer clusters such as Triolith, is of uttermost importance.

PETROS SOUVATZIS

Petros Souvatzis obtained his PhD at Uppsala University at 2007, and after a Postdoc at Los Alamos National Lab he has now returned to Uppsala and is currently employed as a Scientist at the division of materials theory at department of physics and astronomy, Uppsala University. He is the developer of the Self Consistent Ab Initio Lattice Dynamical (SCAILD) method.

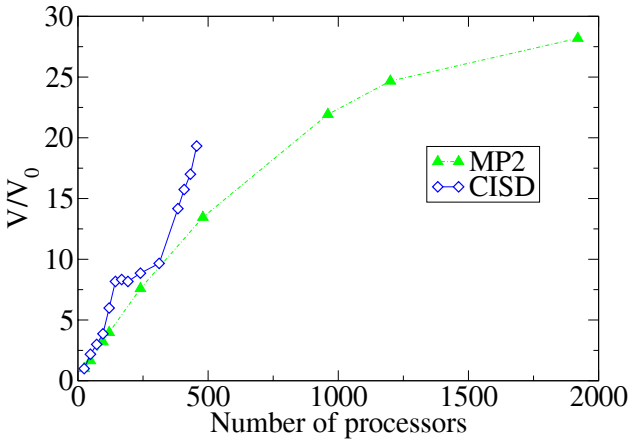



Figure 3: Normalized computational speed as a function of number of processors used, for calculations performed at different levels of theory. Here the green triangles are the computational times of a second order Möller Plesset (MP2) calculation of $H_2O$ using a cc-pVTZ basis set and the empty blue diamonds are the computational times of a singles and doubles configuration interaction (CISD) calculation of $H_2$ using a cc-pVTZ basis set. Here $V_0$ = 1/592 (1/sek) for the MP2 calculation and $V_0$ = 1/425 (1/sek) for the CISD calculation.

## New staff member

Hi! I'm Prashanth and I hail from Bangalore, India. After five years with the Centre for Development of Advanced Computing, Pune India, as a grid and HPC systems administrator, I joined NSC as a system expert, in January 2013. At NSC, I'll be responsible for ushering in and maintaining NSC-LIU's Earth Sciences Grid Federation data node. Apart from my experience of working with buggy grid middleware (which I feel will come very handy here!), I'm proficient in C, Python and Bash programming. I'm very excited to be a part of NSC and I look forward to the challenges ahead!

PRASHANTH DWARAKANATH, NSC

## Better stability and performance for graphical applications on Triolith

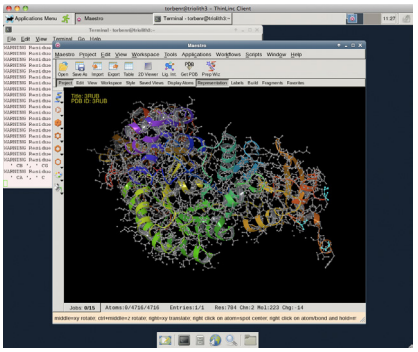
At NSC, we are currently running a pilot project with a remote desktop solution from Cendio AB. The solution is called ThinLinc and only requires that you perform a very simple installation of the ThinLinc client software on your local machine. The client software can then be used to connect to Triolith and get a complete, but so far minimalistic desktop environment based on Xfce. Within this desktop environment, you can launch graphical applications in the usual manner from a terminal command line with 'module load modulename' followed by the name of the binary.

ThinLinc is based on VNC technology for graphics transfer, so running graphical applications in ThinLinc will generally lead to significantly higher performance and will be more stable compared with using a local X-server and accessing Triolith with X-forwarding. Especially applications that use OpenGL for 3D graphics will run with higher performance in ThinLinc, because ThinLinc also supports the use of VirtualGL to provide server-side hardware acceleration of OpenGL. Hence, applications such as GaussView, VMD, Maestro, and VESTA run with better performance using ThinLinc.

The ThinLinc pilot project runs until the end of April this year, but NSC intents to set up a more permanent system for providing good and stable performance of graphical applications, based on either ThinLinc or another product with similar capabilities.

The ThinLinc client software is free and is available for Windows, Mac OS X, Linux, and Solaris. For information regarding how to use ThinLinc to access Triolith, please read: http://www.nsc.liu.se/systems/triolith/triolith_thinlinc_pilottest_guide.html

TORBEN RASMUSSEN, NSC



## NSC continues with iRODS

iRODS is a policy-based Data Management for managing community driven collection life cycles (www.irods.org). Last year NSC arranged a popular iRODS workshop in September. Thanks to the great speakers Leesa Brieger, from RENCI-UNC and Reagan Moore from DICE-UNC the workshop was fully subscribed and prospective participants had unfortunately to be turned down. There were 34 participants from 13 countries representing PRACE as well as ECDS, SND, BILS, SMHI, UPPNEX and CSC attending the workshop.

iRODS is of strategic importance to NSC. Expert knowledge has been built up together with the SNIC SweStore initiative and NSC partner SMHI in a series of iRODS test installations during late 2012 and 2013. Production will start with an BILS iRODS zone after the summer and other production zones will follow later 2013.

Participants at the iRODS workshop at NSC in September 26–28, 2012

At NSC we are looking forward to an exciting future with iRODS. The agenda and presentations from the workshop can be found at the workshop homepage http://www.nsc.liu.se/irods2012/presentations.html.

TOM LANGBORG, NSC

# Triolith, Sweden's Fastest Supercomputer

Triolith has now been in full production since November 2012. The first parts arrived at NSC in June and was taken into production during July. With only a fifth of the total size installed, Triolith already outperformed Neolith. The bulk of compute servers arrived after the summer and during the acceptance period we manage to trim the Linpack number for the november-list of Top500 to a respectable 303 Teraflop/s which positioned Triolith as number 83 on the list of world fastest supercomputers.

The inauguration took place on January 24 at the Linköping University Club. Triolith was inaugurated by the Linköping University Rector Helen Dannetun in the presence of the SNIC Director Jacko Koster, representatives from the manufacturer HP, vendor GoVirtual and a crowd of more than 50 Triolith users.

Triolith consist of 1200 HP SL230s Gen8 servers with dual Intel E5-2660 processors connected together with an Mellanox FDR Infiniband. It has 42.5 Terabyte of memory in total and the aggregated memory bandwidth is about 89 Terabyte/s. The aggregated theoretical performance at nominal clock frequency (2.2 GHz) is 338 Teraflop/s. Since the cores in the latest Intel Xeon processors are eager to increase their clock-frequency over the nominal even at normal conditions, it is somewhat difficult to define "peak" performance. It depends on the thermal conditions for each core.

Triolith was delivered and installed by HP and GoVirtual Nordic AB. Continuing NSC's tradition, we develop and install our own maintained stack of system software on Triolith. This time we made a major overhaul of the system that provide the servers with the kernel and disk image as the old framework was a little fragile during boot. The new system use, as before, BitTorrent to distribute the disk image but instead of performing a kexec to bootstrap into the final production kernel, we now populate the kernel using Dracut.

During summertime this year, right after the new data center building is ready, Triolith will move to the new computer room. At the same time the system will be expanded from 1200 to 1520 compute servers. In general, we are very confident of the overall stability and performance of Triolith and we are certain it will contribute to many research results during its lifetime.

NICLAS ANDERSSON,
TECHNICAL DIRECTOR, NSC

Linköping University

National Supercomputer Centre, Linköpings universitet, 581 83 Linköping, Sweden
Tel: +46 (0) 13-28 26 18, fax: +46 (0) 13-28 25 35, e-mail: info@nsc.liu.se
www.nsc.liu.se

Editor: Peter Münger
Use NSC Express, www.nsc.liu.se/express, or send an e-mail to support@nsc.liu.se to subscribe for a paper or electronic version.