Enclosure 1b. Category 1 Application form 2015 – English version

APPLICATIONS ARE PREFERABLY DRAWN UP IN ENGLISH. AN ENGLISH TRANSLATION HAS TO BE ENCLOSED WITH APPLICATIONS SUBMITTED IN DUTCH.

The application form is available in Dutch on the website https://vscentrum.be/.

Title of the application:

Assessing the accuracy of a screened hybrid functional for property predictions of elemental solids

Name and first name of the applicant:

Lejaeghere Kurt

Institution:

Ghent University

Research group / department:

EA17 / Center for Molecular Modeling

Title / position:

BOF fellow (postdoctoral researcher)

email address:

Kurt.Lejaeghere@UGent.be

Total computing time that is needed, in node days:

1742

Total disk storage that is applied for (in GiB):

163 GB SCRATCH; 12 GB DATA

1. Title of the research project (with IWETO or FRIS link if available) within the framework of which computing time is applied for:

Direction- and temperature-dependent properties of disordered crystals from first principles

2. Describe your research project in short. Explicitly mention the scientific questions that you are planning to address and the overall scientific goals of the project. (max. 1 A4 in Arial 12):

Density-functional theory (DFT) is a quantum physical method that is used increasingly often to design new materials rather than provide a posteriori explanations for observed materials behaviour. Using computational methods in such a way necessitates a thorough understanding of the quality of the predicted values, however. We recently published an exhaustive study of the agreement between DFT in the widely used PBE approximation and experiment for 71 elements (DOI: 10.1080/10408436.2013.772503). This highly cited work showed some critical shortcomings of the approximation, suggesting DFT-PBE to provide predictions of an overall sufficient quality, but to fail for some particular cases. This is mainly due to the (semi)local character of the used PBE functional. Nevertheless, for the research project on disordered crystals, it is essential to obtain reliable values for these special cases as well. This will then allow us to study the elastic properties of a large range of disordered tungsten crystals for nuclear fusion applications.

The current proposal therefore aims to go beyond the popular PBE approach. We will perform a similar error quantification for DFT in the HSE06 approximation as we did for PBE. The HSE06 functional (DOI: 10.1063/1.1564060 and 10.1063/1.220459) combines semilocal contributions with nonlocal exact exchange up to a certain screening distance. It therefore offers a much better representation of the exchange potential, yielding more reliable results. Its agreement with experiment has not been thoroughly quantified yet, however, nor has its expected superior performance been confirmed for all elements or has the sensitivity to the starting potential been checked. This TIER1 proposal therefore aims to perform such study. This effort is performed in collaboration with the Computational Materials Physics group of the University of Vienna, which developed the VASP code and included the HSE06 functional in it.

3. Provide an abstract (10 lines) for scientific communication on the website in layman's terms. See also item 13 of this application form.

Over the last decades, density-functional theory (DFT) has become the work horse of computational materials science. Nevertheless, for several critical applications the most popular approximations do not suffice. The first step towards more accurate predictions lies in the inclusion of nonlocal effects. In this respect, the method by Heyd et al., called HSE06, was found to yield good results for solids. This project therefore aims to quantify and analyse the obtainable accuracy of this computationally more intensive approach and compare to low-cost DFT methods for all elements.

4. Financing institution or channel, financing the research project in full or in part (FWO, BOF, IWT, EU, ...): Please attach the confirmation letter as enclosure (see instructions in enclosure 4 "EasyChair proposals submission procedure").

BOF (confirmation letter at the end of this document)

5. Name and email address of the promoter(s) of the research project:

Prof. Dr. ir. Veronique Van Speybroeck

Prof. Dr. Stefaan Cottenier

6. Billing address to which the payment invoice will be sent to:

Kurt Lejaeghere
Center for Molecular Modeling
Technologiepark 903
9052 Zwijnaarde

7. Persons mandated by the Applicant to compute on the Tier1 within the framework of the present project: Please provide for every person:
   - name and first name
   - institution
   - research group / department
   - title / position
   - experience of using HPC resources in the past (Tier0/Tier1/Tier2 infrastructure in Belgium and abroad)

Lejaeghere Kurt

Ghent University
EA17 / Center for Molecular Modeling
BOF postdoctoral fellow
experience with local CMM clusters, local clusters of the Computational Materials Physics Group of the University of Vienna, HPC UGent TIER2 clusters and the TIER1 machine muk (vsc40323)

Van Speybroeck Veronique

Ghent University
EA17 / Center for Molecular Modeling
Full professor
experience with local CMM clusters and HPC UGent TIER2 clusters (vsc40021)

Cottenier Stefaan

Ghent University
EA10 / WE05 / Center for Molecular Modeling
Assistant professor
experience with local CMM clusters and with Leuven and HPC UGent TIER2 clusters (vsc40026)

8. Explain why this project needs to run on a Tier1 system, why the machine you have requested is suitable for the project and how the use of the system will enable the science proposed (max. ½ A4 in Arial 12).

Although each individual VASP calculation can be performed on TIER2 infrastructure, the combination of multinode jobs and large job numbers make it very cumbersome to deal with the long waiting times on TIER2 infrastructure. Moreover, even when using 20 16-core nodes non-stop (which is currently far from possible on the UGent TIER2 clusters), 90 days of wall-clock time would be required. Finally, the proposed jobs scale favourably for small numbers of nodes (see question 10), allowing us to benefit from a limited multinode parallelization (larger systems in the set can be parallelized further). Hence, only the TIER1 machine allows us to complete such a computational task efficiently.

9. Justify the number of node days requested. This should include information such as: number and nature of computing tasks, software used, and the sequence in which they will be performed. Indicate for each typical computing task the required resources:

    - wall clock time (note that 3 days is the maximal wall clock time for any job; checkpointing should be used for longer run times)
    - memory (maximum 64 GiB/node)
    - number of nodes
    - number of CPU cores
    - disk space (estimated volume in GiB and the total number of files); make a clear distinction between usage of Tier2 DATA/HOME partitions and the Tier1 SCRATCH partition
    - number of tasks, and an indication of how many such tasks would be submitted concurrently.

    This information should take the form of a table (an example is provided as Table 2 in the appendix). Provide additional descriptions of the computing tasks and comments as needed. Resource estimates should be preferably based on the results of actual calculations on Tier1 (via, e.g., a Starting Grant) for system/problem sizes that are on par with those of the intended computing tasks (e.g., same mesh sizes, actual molecular system, …). If not, provide the name, architecture, #cores, memory, etc. of the machine that was used to obtain these results and explain how you have calculated/rescaled the wall clock times, number of cores, etc.

    (max. 1 A4 Arial 12).

A typical HSE optimization takes several steps, which need to be repeated for each volume of an equation of state (EOS) (13 per material). Moreover, the most accurate results are obtained when the EOS is reinitialized around the correct volume. This entails the following work flow: PBE and HSE (low-precision) preconvergence for an EOS around a predefined volume; PBE and HSE (low-precision) preconvergence around the correct equilibrium volume; data collection at normal precision; precision check for 1 scf step at high precision. This protocol is followed for 31 cubic materials (no relaxations needed, class A) and 40 non-cubic ones (relaxations needed, class B). Finally, the results are heavily influenced by the used

PAW potential. We therefore assess 3 sets of potentials (2 PBE ones and 1 LDA).

Timing and memory requirements were obtained from benchmarking on delcatty (UGent TIER2 cluster with the same architecture as the TIER1).

class A (3 x 13 x 31 tasks)	*Computing time*	*Memory*	*# Nodes*	*Disk space (# files)*	
				SCRATCH	*DATA*
GGA	20 core minutes	2 GB	1	8 MB (18+1)	0.5 MB (1)
HSE low precision	8 core hours	8 GB	1	8 MB (18+1)	0.5 MB (1)
GGA	20 core minutes	2 GB	1	8 MB (18+1)	0.5 MB (1)
HSE low precision	8 core hours	8 GB	1	8 MB (18+1)	0.5 MB (1)
HSE normal precision	30 core hours	8 GB	≈4	12 MB (18+1)	0.5 MB (1)
HSE high precision (1 scf step)	46 core hours	8 GB	≈4	12 MB (18+1)	0.5 MB (1)
Subtotal	**5.8 node hours**	**8 GB**	**1-4**	**56 MB (18+6)**	**3 MB (6)**
class B (3 x 13 x 40 tasks)	*Computing time*	*Memory*	*# Nodes*	*Disk space*	
				SCRATCH	*DATA*
GGA	20 core minutes	2 GB	1	8 MB (18+1)	0.5 MB (1)
HSE low precision	10 relaxation steps x 8 core hours	8 GB	1	10 MB (18+1)	1.5 MB (1)
GGA	20 core minutes	2 GB	1	8 MB (18+1)	0.5 MB (1)
HSE low precision	10 relaxation steps x 8 core hours	8 GB	1	10 MB (18+1)	1.5 MB (1)
HSE normal precision	5 relaxation steps x 30 core hours	8 GB	≈4	13 MB (18+1)	1.0 MB (1)
HSE high precision (1 scf step)	46 core hours	8 GB	≈4	12 MB (18+1)	0.5 MB (1)
Subtotal	**22.3 node hours**	**8 GB**	**1-4**	**61 MB (18+6)**	**5.5 MB (6)**
Total	**1742 node days**	**8 GB**	**1-4**	**163 GB (66 456)**	**12 GB (16 614)**

10. Describe the software required to perform the computing task(s). Please clearly provide the following per item in this regard:

    - a reference to the software's web page
    - the software license system (open source, GPL, etc.)

- if there is no free academic use of the software, state which license makes the installation and the use valid on the Tier1 by the Applicant (+ add a copy of the signed license)
- if need be, which license server will be used (name + IP address)
- whether the software is already available on the Tier1 (see https://vscentrum.be/nl/Tier1-rekenen) and, if this is not the case, compilation and installation instructions (possibly with reference to existing Tier2 installation)

VASP

http://www.vasp.at/
non-exclusive academic licence (see attachment)
available on TIER1
MPI is used for parallellization, no vSMP will be used

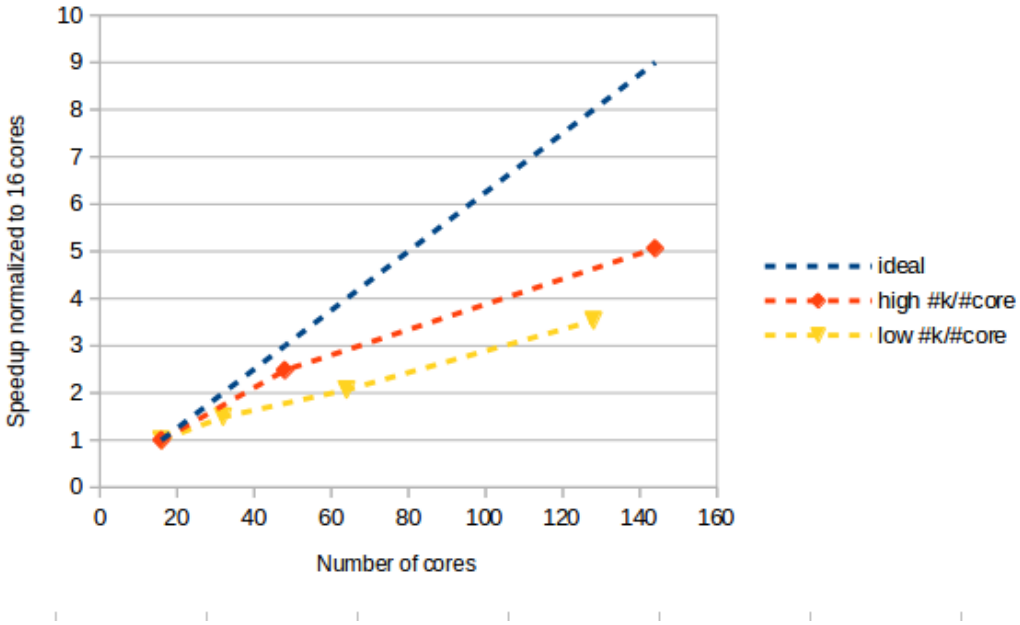
ASE

https://wiki.fysik.dtu.dk/ase/
GNU Lesser General Public Licence v2.1+
available on TIER1

Provide the results of scaling tests that were conducted with this software, preferably on Tier1 (using, e.g., a Starting Grant) for system/problem sizes that are on par with those of the intended computing tasks (e.g., same mesh sizes, actual molecular system, …). If not, provide the name, architecture, #cores, memory, etc. of the machine that was used to obtain these results.
Provide both a table and scaling plot such as table 1 and plot 1 in the appendix (max. 2 A4 in Arial 12).

All scaling tests were performed on delcatty (UGent TIER2 cluster with the same architecture as the TIER1) and apply to a normal-accuracy HSE calculation for silver. We note a distinct improvement in scaling when k-point parallelization is optimized (bold typesetting, in the benchmark case there were 72 k-points). This indicates an acceptable parallelization behaviour for limited numbers of nodes, *as long as the number of cores is chosen compatible* with the investigated material. We therefore aim to limit ourselves to calculations of up to 3-4 nodes per job with an optimal k-point/core ratio.

# nodes	# cores	absolute timing (s)	speedup	# cores x timing	# parallel k-points / # cores
1	**16**	**1 751**	**1.0000**	**28 016**	**0.5**
2	32	1 175	1.4902	37 600	0.25
3	**48**	**706**	**2.4802**	**33 888**	**0.5**
4	64	848	2.0649	54 272	0.125
8	128	496	3.5302	63 488	0.0625
9	**144**	**346**	**5.0607**	**49 824**	**0.5**



11. Describe how you will manage the resources requested in the period during which the task is to be performed. What usage pattern do you anticipate (similar usage on monthly basis, bursts, …)?

The intention of this project is to maintain a constant flow of jobs, unless computational problems arise (convergence issues, …).

12. List the granted computing time allocations to the promoter(s) of this research project, on the Flemish Tier1 system, as well as other

Tier1 and Tier0 systems. Also, describe the scientific output obtained within the framework of computing time that was granted during the past two years on the Flemish Tier1 or on other Tier1 or Tier0 supercomputers. DOI links are sufficient.

"High-throughput screenen van ternaire wolfraamlegeringen met DFT", pilot phase TIER1 project, 4053 node days granted to Kurt Lejaeghere (17-07-2013 to 31-12-2013).

"High-throughput screening of ternary tungsten alloys with DFT", TIER1 project, 4000 node days granted to Kurt Lejaeghere (04-07-2014 to 31-12-2014).

Results from this previous TIER1 work have not been published in a journal yet, but have appeared partly in the PhD thesis of Kurt Lejaeghere (ISBN 978-90-8578-690-0) and in posters and talks at conferences (EMRS spring meeting (Lille, May 26-30 2014); Frontiers of first-principles simulations: materials design and discovery (Berlin, February 1-5 2015)). The two TIER1 projects have allowed us to assemble most of the necessary data. Only for 25 materials, some manual adjustments need to be made (on TIER2 infrastructure). A publication is planned when these calculations have been completed.

13. Are the applicants of this application bound by a confidentiality agreement? If so, the title and the abstract of this application will not be published on the website of the Hercules Foundation / Flemish Supercomputer Center.

no

Should you have any questions or encounter any difficulties during the electronic submission of an Application, please contact by e-mail:
Associatie KU Leuven: hpcinfo@kuleuven.be
Associatie Universiteit Gent: hpc@ugent.be
Associatie Universiteit Hogescholen Antwerpen: hpc@uantwerpen.be
Associatie Universiteit Hogescholen Limburg: geertjan.bex@uhasselt.be
Universitaire Associatie Brussel: rosette.vandenbroucke@vub.ac.be
For the other institutions: marc.luwel@herculesstichting.be

Appendix: Example tables and plots

Table 1

# nodes	# cores	absolute timing (s)	speedup	# cores x timing
16	256	189.6	1.0000	
32	512	99.0	1.9154	
64	1024	55.6	3.4088	
128	2048	30.8	6.1376	

Plot 1

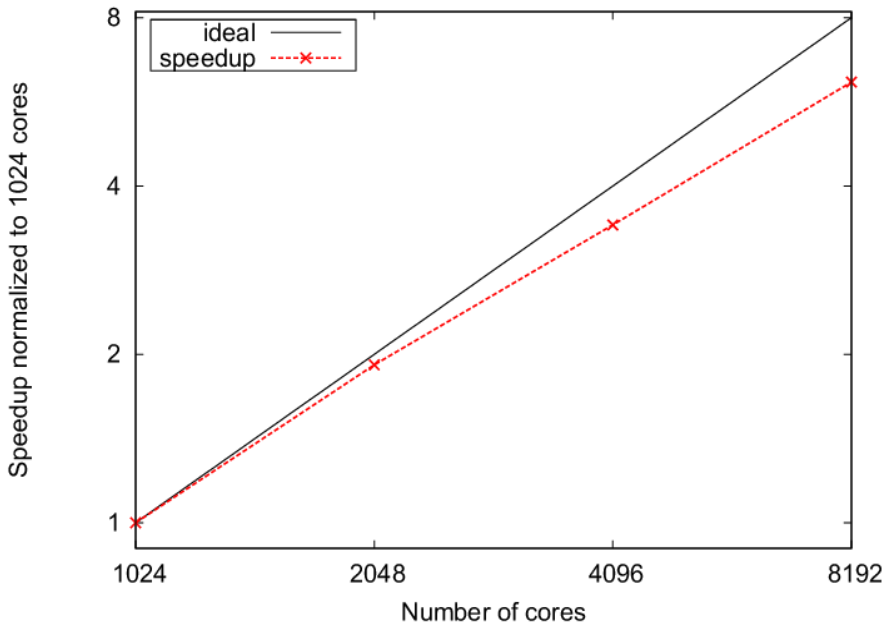


Table 2

	Node day calculation							Storage volume estimate	
Computational task	# of such tasks	Wall clock time (days) per task	# Tier1 nodes per task	# node days per task	# CPU cores per task	Memory usage (GiB) / node per task	OpenMP / MPI / hybrid / vSMP	Tier2 DATA/HOME volume (GiB) + number of files	Tier1 SCRATCH volume (GiB) + number of files
Task1	A	B	C	= A x B x C					
Task2	A	B	C	= A x B x C					

Aanstelling tot voltijds doctor-assistent onderzoeksmandaat van de heer Kurt Lejaeghere

DE RECTOR VAN DE UNIVERSITEIT GENT

Gelet op het bijzonder decreet van 26 juni 1991 betreffende de Universiteit Gent en het Universitair Centrum Antwerpen;

Gelet op het decreet van 12 juni 1991 betreffende de universiteiten in de Vlaamse Gemeenschap, zoals het is gewijzigd, inzonderheid op hoofdstuk IV;

Gelet op het besluit van de Vlaamse Regering van 4 mei 2001 tot vaststelling van de salarisschalen van het assisterend academisch personeel van de universiteiten van de Vlaamse Gemeenschap, zoals gewijzigd;

Gelet op het reglement met betrekking tot de mandaten van doctor-assistent met hoofdzakelijk onderzoeksopdracht (bestuurscollege 3 juli 2008, gewijzigd 23 juni 2011);

Gelet op het besluit van de raad van bestuur van de Universiteit Gent van 8 november 2013 tot vaststelling van de personeelsformatie bij de Universiteit Gent;

Gelet op het voorstel van de onderzoeksraad van 1 juli 2014 tot toekenning aan de heer Kurt Lejaeghere van een mandaat als doctor-assistent met hoofdzakelijk een onderzoeksopdracht voor drie jaar;

Overwegende dat de heer Kurt Lejaeghere voldoet aan de gestelde functie-eisen;

Onverminderd de prerogatieven van de regeringscommissaris en van de afgevaardigde van Financiën zoals die zijn vastgelegd in het bovengenoemde decreet van 12 juni 1991, inzonderheid in de artikelen 177 en 180 voor wat de uitvoering van de beslissingen betreft;

BESLUIT:

Artikel 1.- De heer Kurt Lejaeghere (geboren op 28.03.1987), wordt met ingang van 1 oktober 2014 voor drie jaar aangesteld tot voltijds doctor-assistent onderzoeksmandaat bij de vakgroep Toegepaste Fysica van de Universiteit Gent.

De heer Kurt Lejaeghere geniet een salaris berekend op basis van de salarisschaal AAP5 (salaris à 100%: min. 29.069,73 EUR – max. 45.317,25 EUR, thans uitbetaald à 160,84%), rekening houdend met eventueel verworven anciënniteit, ten laste van de werkingsuitkeringen van de Universiteit.

Artikel 2.- Dit besluit zal aan de betrokkene medegedeeld worden.

Gent,

15 SEP. 2014

Prof. dr. A. De Paepe
Rector