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:

Computational Discovery of Quaternary Zintl Phases

Name and first name of the applicant:

Titus Crepain

Institution:

Ghent University

Research group / department:

Center for Molecular Modeling (CMM)

Title / position:

Master thesis student

email address:

titus.crepain@ugent.be

stefaan.cottenier@ugent.be

Total computing time that is needed, in node days:

3076

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

3500 GiB

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

FWO (G.0402.11N)

"Mercury, the Earth, and exoplanets: the crystal structure and other properties of the inner core of terrestrial planets by Density Functional Theory and a genetic search algorithm."

(TIER1-time will be paid from this project, as long as no dedicated funding is achieved for materials discovery, which is the actual topic of this proposal.)

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):

A large majority of all possibly existing binary solids have effectively been discovered and described. For ternary solids, the estimated number of unknown ones is still larger than the number of known ones. With quaternary solids, one enters a Terra Incognita where only a negligible fraction of all possible solids has been described. This is nicely illustrated by the so-called Zintl phases – crystals formed by combining s- and early p-elements, showing a chemical bond that has characteristics of ionic as well as of covalent bonding. It is a combinatorial exercise to find that 5600 different quartets can be formed when picking one element from the s-groups of the periodic table, and three elements from the early p-groups (5600= $10 \binom{16}{3}$). For each quartet, several plausible valence-balanced stoichiometries can be suggested, and for each stoichiometry several (as yet undetermined) plausible crystal structures are possible. Even with conservative estimates for the latter two degrees of freedom, one rapidly ends up with a search space that contains several hundreds of thousands of potentially existing quaternary Zintl phases (QZP).

How many QZP are experimentally known? According to the Inorganic Crystal Structure Database, which contains 160.000 experimentally known crystals, as few as 50 can be labeled as a true QZP. Only a couple of new QZP are discovered each year.

We demonstrated by a recent modest 1000 node day TIER1 project that it is possible to speed up dramatically the discovery rate of new QZP. To achieve this, we took a crystal structure that is 'common' for QZP (i.e. 6 QZP are known to adopt this crystal structure), and decorated the positions with several thousands of quartets of elements. The stability of all of these crystals was subsequently determined by quantum simulations (density-functional theory). This resulted in a list of crystals that are energetically stable and which should therefore exist. These can now become the prime targets of experimental researchers who try to synthesize new QZP.

Building on this success, we are now proposing to tackle a new family that is even richer: the 1:1:2:2 stoichiometry, for which a few experimentally existing cases are known in the Ibam crystal structure (e.g. $NaInK_2P_2$). Not only will this very likely once again substantially increase the list of relevant targets for experimental synthesis, but having exhaustive data across multiple stoichiometries and crystal structures will gradually allow to mine these data for heuristic rules to predict QZP-formation.

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

There is a large group of less-known solids that share properties with the ionic solids as much as with the covalent solids: Zintl phases. Scientists have reasons to expect interesting and useful materials among the Zintl phases built from four (or more) elements – the quaternary Zintl phases (QZP). The 50 true QZP that are experimentally known are negligible compared to the hundreds of thousands of QZP that might potentially exist. We have recently demonstrated that computational screening can dramatically speed up the discovery rate of new QZP. In this project, we want screen for stable QZP with the 1:1:2:2 stoichiometry, and start constructing a database that can be mined to understand QZP-formation.

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").

FWO

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

Prof. Dr. Stefaan Cottenier (stefaan.cottenier@ugent.be)

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

Stefaan Cottenier
t.a.v. Wim Dewitte
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)

   Crepain Titus
   Ghent University
   WE05 / Center for Molecular Modeling
   Master thesis student
   Experience with the HPC UGent TIER2 clusters, which he acquired during the previous academic year in the context of courses, and which he will deepen during the coming summer when he will get an intensive training in preparation of his thesis work. In order to guarantee efficient use of TIER1-time, the expert users involved in this project will take care of the TIER1-aspects during the first half of this project, and Titus will gradually take over as his experience grows.

   ir. Sluydts Michael
   Ghent University
   Center for Molecular Modeling
   PhD Fellow
   Extensive experience using the VASP and wien2k ab initio packages on both TIER2 and TIER1 (vsc40479)

Dr. ir. Lejaeghere Kurt
Ghent University
Center for Molecular Modeling
BOF postdoctoral fellow
Extensive 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)

Prof. Dr. Cottenier Stefaan
Ghent University
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).

The proposed work consists of 3456 similar jobs, which require 0.86 node days each. The work flow of such a job has been thoroughly tested on TIER2 and in a previous exploratory TIER1 proposal (1000 node days) . We are now ready to take the next step and address a richer part of the search space. Due to the relatively large number of jobs, this is not feasible on TIER2. The proposed work will be the core of the master thesis of Titus Crepain (to be defended in June 2016), and of the publication based thereupon.

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).

	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
initial volume optimization	3456	0.03	1	103.7	16	15	MPI	0.0 / 2 (zip)	0.3 / 20
Full geometry optimization	3456	0.69	1	2384.6	16	15	MPI	0.0 / 2 (zip)	0.3 / 20
Equation of state generation	3456	0.11	1	380.2	16	15	MPI	0.0 / 14 (zip)	2.0 / 140
Final single point calculation	3456	0.02	1	69.1	16	20	MPI	0.3 / 2 (zip)	0.3 / 20
Density Of States	3456	0.04	1	138.2	16	20	MPI	0.7 / 2 (zip)	0.7 / 20
total				**3075.8**				**3500 / 76000 (zip)** (all 3456 tasks)	**360 / 22000** (peak for 100 simultaneous tasks)

- **initial volume optimization :** the unit cell of the crystal is decorated with the chosen quartet of elements, and without altering the relative positions of the atoms a first estimate of the volume of the unit cell is searched.
- **full geometry optimization :** for the unit cell volume achieved in the previous step, the position of the atoms and the shape of the unit cell is optimized. This is by far the most time-consuming step.
- **equation of state generation :** for the optimized cell resulting from the previous step, an equation of state is determined that yields an accurate volume and bulk modulus.
- **single point calculation :** the crystal is now fully optimized, and an accurate total energy can be determined. This is the number that will be used for the stability assessment.
- **density of states :** a lot of useful information is implicitly contained in the Density of States of the crystal. By computing and storing this information now, relations between stability and macroscopic properties of the crystal can be established later on.

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

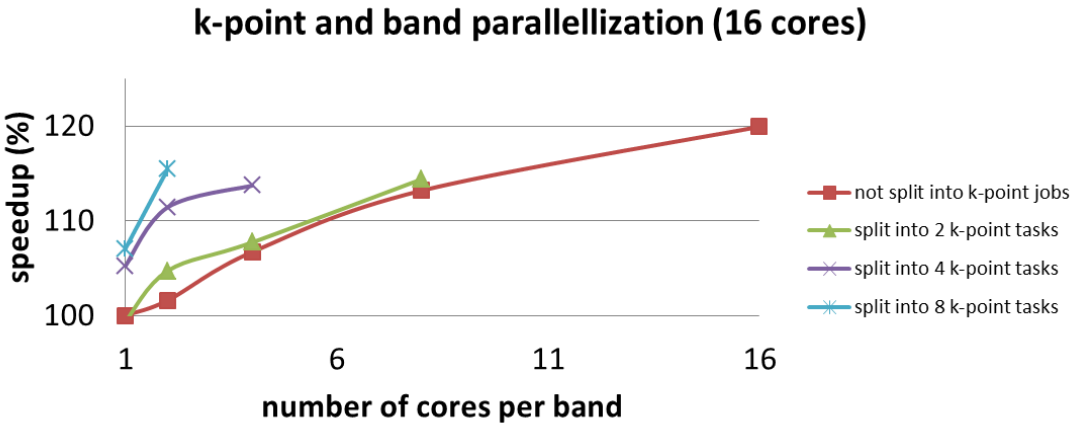
    - a reference to the software's web page
    - the software license system (open source, GPL, etc.)
    - if there is no free academic use of the software, state which license makes the installation and the use valid on the Tier1 by the Applicant (+ add a copy of the signed license)
    - if need be, 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)

    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).

    VASP
    http://www.vasp.at/
    non-exclusive academic licence (see attachment)
    available on TIER1

    A lot of experience about the scaling behavior of VASP on MUK is available, for various types of jobs and levels of parallelization. The picture underneath shows a scaling test for a typical calculation using a QZP with 1:1:1:4 stoichiometry as dealt with in our previous TIER1-project (for the 1:1:2:2 stoichiometry the execution time slightly differs, but the scaling is identical). In this test, two types of parallelization are combined with each other, in different ways: k-point parallelization (=distributing the job into almost independent parts) and band-parallelization (=assigning multiple cores to each electron band, for a given k-point). This can be done in different combinations (see picture) – in every case, always all 16 cores of a

node are used. This test shows that the shortest wall-time is achieved for band-parallelization with 16 cores collaborating on the same electron band, without any k-point parallelization. In this way, it is possible to squeeze 20% more performance out of the node. As the major challenge in this proposal lies in the large number of short-duration jobs (0.86 node-days per job, 3456 jobs needed), it is not useful in this case to parallelize more.



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, …)?

Similar usage on a monthly basis. The actual start of the project would ideally be September 1, with the default duration (6 months).

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.

Finished projects :

Hunting for new quaternary Zintl phases of the 1:1:1:4 (1000 node days)
K. Dumon, M. Sluydts, K. Lejaeghere,S. Cottenier

The present proposal is a direct successor of the abovementioned successfully finished exploratory project. Formation energies for more than 4000 QZP with the 1:1:1:4 stoichiometry have been obtained. Manual inspection of the most promising cases readily provided 10 examples that are stable against any decomposition in binary or ternary phases. We are now automating the analysis process (for which TIER1 is not required), and estimate that we will end up with several hundreds of stable QZP candidates – a number that is larger than the total number of QZP currently known.

This work has been extensively described in the master thesis of Karel Dumon, which will be defended in June 2015. A publication will follow.

High-throughput determination of vacancy trapping enthalpies for the improvement of electronic device production (4420 node days)
M. Sluydts, D.E.P. Vanpoucke

Formation energies for most elements of the periodic table as impurity at several lattice sites in the leading semiconductors silicon and germanium have been calculated. This extensive set of data has been analyzed in terms of prevalence of lattice sites for specific elements, and in terms of vacancy gettering for the production of high-quality silicon and germanium single crystals. A paper is currently under review.

High-throughput screening of ternary tungsten alloys with DFT (4000 node days)
K. Lejaeghere
+
High-throughput screenen van ternaire wolfraamlegeringen met DFT
(4053 node days, pilot phase TIER1 project)
Kurt Lejaeghere

These projects aimed to computationally screen a large set of ternary tungsten alloys for their suitability as first-wall materials in nuclear fusion reactors. Properties of interest were, among other things, melting temperature, ductility and hardness. 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 4 materials, some manual adjustments need to be made (on TIER2 infrastructure). A publication is planned when these calculations have been completed.

Ongoing projects

Ab initio prediction of acoustic anisotropy of Fe, Ni, and FeNi in the Earth's inner core (3106 node days)
J. Jaeken

The heart of this project is a set of crystals with varying concentrations of iron and nickel, and with several different stacking faults. These sample the material that is present at the inner core of the Earth, under high pressure and at high temperature. Density Functional Theory is being used to assess the elastic properties of these materials as a function of pressure. One of the properties that can be derived from this information is the anisotropy of the propagation speed of acoustic waves, which can be directly related to seismographic information.

Assessing the accuracy of a screened hybrid functional for property predictions of elemental solids (1742 node days)
K. Lejaeghere

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.

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

SOFTWARE LICENSE AGREEMENT FOR THE USE OF VASP5.2 BY ACADEMIC INSTITUTIONS

The Universität Wien, Austria (UW in the following) and
Ghent University, Belgium (UG in the following) [1]
conclude the following agreement:

(1) The UG acquires a non-exclusive academic license for the use of the software-package VASP (Vienna ab-initio simulationprogram) for ab-initio local-density-functional total-energy and molecular-dynamics calculations, versions VASP5.2 and VASP4.6, by the research group Functional Nanomaterials (FUNNANO)[2]. Under this licence the use of the software is restricted to a maximum of six researchers or students, all belonging to this research group and to the same organisatorial unit and working at the same location. The licence does not cover the use of VASP by external collaborators working at other institutions.

(2) The license covers access to the source-code, the program documentation and to the data-base for ultrasoft pseudopotentials and PAW-potentials. UW reserves the exclusive property of the software. It declines any liability for the software and any responsibility for the results of calculations produced with the program. The license does not cover any maintenance service for the software or support for its implementation.

(3) The license is not transferable to another research group of UG without the written agreement of UW. UW reserves the right to refuse authorization of such a transfer. A transfer to a research group not belonging to UG is excluded.

(4)The UG guarantees that the software or parts thereof shall not be made accessible to third parties without the explicit written consent of UW. Access to the code and to the data-base shall be made available through an account of the UW. The UG guarantees that the password for this account will be known only to one contact-person and shall not be communicated to temporary co-workers or guests. All installations of the source code, the executable or the data-base must be copy-protected and accessible only to the authorized users.

[1] Please insert here the name of the institution concluding this agreement with UW. This institution must be a legal person and the agreement must be signed by an authorized representative of this institution. Define the acronym (replacing .....) under which this institution is referred to in the text of the agreement.

[2] Please insert here the name and affiliation research group for which the license is acquired

SOFTWARE LICENSE AGREEMENT FOR THE USE OF VASP5.2 BY ACADEMIC INSTITUTIONS

(5) If VASP is used as the basis of further methodological or software-development, UG agrees to make these additions available to UW. UW will also be entitled to include these additions in further releases of VASP.

(6) In future publications of work performed using VASP, the use of the software shall be properly acknowledged, e.g. in the form
"The calculations have been performed using the ab-initio total-energy and molecular-dynamics program VASP (Vienna ab-initio simulation program) developed at the Institut für Materialphysik of the Universität Wien [1-3]."
[1] G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11 169 (1996).
If the PAW-version is used, reference will be made to
[2] G. Kresse and D. Joubert, Phys. Rev. **59**, 1758 (1999).
If special features implemented in VASP will have been used, reference should be made to the relevant publications as listed on the VASP home-page.

(7) The UG accepts to pay to UW a licence fee Euro 4.000,- (fourthousand Euro). The licence fee is strongly discounted and applies only to academic institutions with undergraduate teaching.

(8) The licensee will use VASP exclusively for non-profit research. If VASP is used in contractual research in cooperation with or for industry or for military institutions, the financial conditions will have to be re-negotiated.

(9) UW declares that it has the full power and authority to grant the rights granted in this agreement without the consent of any other person, and that the license and use of the software by the licensee will not in any way constitute an infringement or other violation of any copyright, proprietory right or any other rights of any third party.

(10) Any disputes arising from the license agreement are subject to the laws of the Republic of Austria.

(11) The terms of this agreement shall prevail any terms or conditions of the licensee.

SOFTWARE LICENSE AGREEMENT FOR THE USE OF VASP5.2 BY ACADEMIC INSTITUTIONS

For the Universität Wien:

Jürgen Hafner
Fakultät für Physik, Universität Wien
Sensengasse 8/12, A-1090 Wien, Austria

Date

For the UG

Name (in print): Michel Waroquier
Institution: Faculty of Sciences, Ghent University

Address: Technologiepark 903, BE-9052 Zwijnaarde, Belgium

Date: 26 January 2010

For the research group entitled to use VASP5.2:

Name (in print): Veronique Van Speybroeck (FUNNANO)