Enclosure 1b. Category 1 Application form – 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 English on the website
https://www.vscentrum.be/en/access-and-infrastructure/project-access-tier1

Title of the application:

Assessing the accuracy of an efficient meta-GGA 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:

FWO postdoctoral researcher

e-mail address:

Kurt.Lejaeghere@UGent.be

Total computing time that is needed, in node days:

640

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

55 GiB DATA / 200 GiB SCRATCH

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

Unravelling energetic and mechanical properties of framework materials by means of advanced electronic-structure methods

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. Nevertheless, for the research project on framework materials, 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 nanoporous and dense framework materials.

The current proposal therefore aims to go beyond the popular PBE approach. We will perform a similar error quantification for DFT in the SCAN approximation as we did for PBE. The SCAN functional (DOI: 10.1103/PhysRevLett.115.036402 and 10.1038/nchem.2535) was built to adhere to all known constraints for a semilocal exchange-correlation functional and to be appropriately normed for a number of prototypical situations. It was suggested to perform much better than standard GGA functionals, sometimes even outperforming hybrid functionals. 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.

3. Provide an abstract (10 lines) for scientific communication on the website in layman's terms. See also item 12 of this application form.

No abstract of this work may be published on the website (see item 12).

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 3 “EasyChair proposals submission procedure”).

FWO (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

(Veronique.VanSpeybroeck@UGent.be)

Prof. Dr. Stefaan Cottenier

(Stefaan.Cottenier@UGent.be)

6. Persons mandated by the Applicant to compute on the Tier-1 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 (Tier-0/Tier-1/Tier-2 infrastructure in Belgium and abroad)

Lejaeghere Kurt

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

De Vos Arthur

- Ghent University
- EA17 / Center for Molecular Modeling
- FWO doctoral fellow
- experience with local CMM clusters, HPC UGent TIER2 clusters and the TIER1 machine muk (vsc41228)

7. Explain why this project needs to run on a Tier-1 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 large number of jobs (9230) makes it very cumbersome to deal with the long waiting times on TIER2 infrastructure. We plan to group these jobs as much as possible (e.g. batches of 13 jobs at a time), but the number of jobs will still be too high to submit in the often full queues on TIER2 infrastructure. Moreover, even when using 20 16-core nodes non-stop (which is rarely possible on the UGent TIER2 clusters), 56 days of wall-clock time would be required. Hence, only the TIER1 machine allows us to complete such a computational task efficiently and within a time scale that allows us to compete with other high-profile research groups.

8. 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 256 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 Tier-2 DATA/HOME partitions and the Tier-1 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 Tier-1 (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 table with timing and memory requirements is given in the Appendix (Table 3) and is based on VASP test calculations on the copper, rhodium

and arsenic system (single-point calculation) on the intended BrENIAC cluster.

Our test set to assess the accuracy of DFT functionals consists of 71 materials, 31 of which are cubic (task 1 in Appendix Table 3) and 40 of which are noncubic (task 2 in Appendix Table 3). Their equation of state (EOS) is reconstructed by calculating the energy for 13 volumes per material. For cubic materials, only a single-point calculation is needed for each volume (0.004 node days on BrENIAC per material and per volume), while noncubic materials require an optimization which on average takes 30 relaxation steps per volume (30 x 0.004 node days = 0.12 node days per material and per volume). In addition, the results may be heavily influenced by the used PAW potential. In fact, no dedicated potentials are available for the used SCAN functional. To exclude unwanted inaccuracies, we will therefore assess the available potentials, which are based on the LDA, PW91 and PBE functionals. There are 710 potentials available, yielding approximately 10 potentials to be tested per material. This means that in total, 10 x 31 x 13 calculations need to be performed for cubic materials and 10 x 40 x 13 calculations for noncubic materials.

Although the intended calculations correspond to a huge number of jobs (9230), we intend to group calculations as much as possible (e.g. batches of 13 calculations in a single job = 1 material with 1 potential), as not to strain the TIER1 queuing system too much. The ASE Python module will be used to do this. We will moreover spread the jobs over time, never submitting more than 50 (grouped) jobs at once. Finally, the estimated SCRATCH disk space totals 923 GiB, but we will monitor the calculations and remove all restart files as soon as possible, reducing the need for SCRATCH space to 200 GiB at most.

9. 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 Tier-1 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 Tier-1 (see https://www.vscentrum.be/cluster-doc/software/tier1-muk) and, if this is not the case, compilation and installation instructions (possibly with reference to existing Tier-2 installation)

VASP

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

ASE

https://wiki.fysik.dtu.dk/ase/
GNU Lesser General Public Licence v2.1+
available on UGent HPC and Muk, but not on BrENIAC yet
compilation and installation instructions available on ASE website

> Provide the results of scaling tests that were conducted with this software, preferably on Tier-1 (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).

The results of some VASP scaling tests, performed on BrENIAC, are available in Tables 1 and 2 and Plot 1 in the Appendix, as well as in attachment. Tests on 1 node (Table 2) were performed with a varied selection of systems that will be investigated in this proposal (Cu, Rh, As), but at lower computational settings to allow a quick sweep of parameter space. Tests on multiple nodes (Table 1) were reported for a Fe-N metal (most of the materials investigated in this proposal are metals) and showed decent scaling up to at least 7 nodes. We did not perform new multinode scaling tests for the current proposal, since our extended benchmark studies (see attachment) showed a close similarity in scaling behaviour for all systems, despite their large differences, i.e. for both large and smaller systems, and for both metals and nonmetals. In addition, this proposal does not intend to use more than 1 node per calculation, as the computational cost for a single calculation is rather limited. However, we did test the optimal combination of parallelization over k-points or parallelization over electronic bands (Table 1), as our extended benchmark (see attachment) showed the optimal parallelization settings to be highly system-dependent. Our results for Cu, Rh and As show that it is best to use 7 cores per electronic band (NCORE) when 1 k-point is treated at a time. In line with our previous benchmark results, which showed a high k-point parallelization to be favourable, we will perform all calculations with NCORE = 7 and KPAR = 4 on 1 node each.

10. 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, …).

11. List the granted computing time allocations to the promoter(s) of this research project, on the Flemish Tier-1 system, as well as other Tier-1 and Tier-0 systems. Also, describe the scientific output obtained within the framework of computing time that was granted during the past two years on the Flemish Tier-1 or on other Tier-1 or Tier-0 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).

"Assessing the accuracy of a screened hybrid functional for property predictions of elemental solids", TIER1 project, 1742 node days granted to Kurt Lejaeghere (22-03-2015 to 15-11-2015).

"Error quantification for the adiabatic connection fluctuation-dissipation theorem in the random phase approximation (ACFDT-RPA)", TIER1b pilot project, 1000 node days granted to Kurt Lejaeghere (13-07-2016 to 17-10-2016).

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); ICAMM 2016 (Rennes, September 5-7 2016)). A publication is planned in collaboration with the University of Vienna that combines results from calculations on TIER2 infrastructure with the results from the two most recent TIER1 project and the current proposal. A publication on the results of the first two TIER1 projects is currently on hold due to changed priorities, but is expected to be written in 2017.

12. 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 FWO / Flemish Supercomputer Center.

yes

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: caroline.volckaert@FWO.be

Appendix: Tables and plots

Table 1

# nodes	# cores	absolute timing (s)	speedup	# cores x timing (s)
1	28	5601	1.0000	156 826
2	56	3229	1.7348	180 804
4	112	1710	3.2752	191 531
7	196	1031	5.4307	202 142

Plot 1

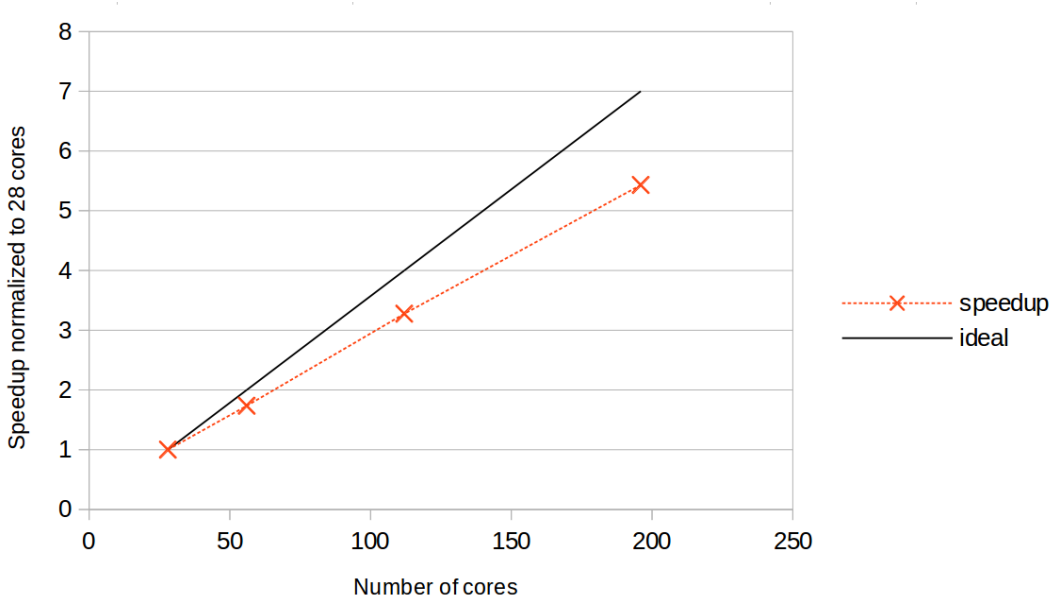


Table 2

	absolute timing on 1 node with KPAR = 1 [s]			
element	NCORE = 1	NCORE = 7	NCORE = 14	NCORE = 28
Cu	61.5	60.5	78.7	138.4
Rh	144.6	113.6	156.1	283.0
As	200.7	161.7	198.5	317.4

Table 3

	Node day calculation							Storage volume estimate	
Computational task	# of such tasks	Wall clock time (days) per task	# Tier-1 nodes per task	# total node days	# CPU cores per task	Memory usage (GiB) / node per task	OpenMP / MPI / hybrid / vSMP	Tier-2 DATA/HOME volume (GiB) + number of files (total)	Tier-1 SCRATCH volume (GiB) + number of files (total)
single-volume calculation cubic material	31 x 13 x 10	0.004	1	16	28	10-20	MPI	5 GiB (56 420 files)	403 GiB (68 510 files)
single-volume optimization noncubic material	40 x 13 x 10	0.12	1	624	28	10-30	MPI	50 GiB (72 800 files)	520 (88 400 files)

fwo
Opening
new
horizons

De heer Kurt Lejaeghere
Beernemsteenweg 42
8750 WINGENE

uw kenmerk

ons kenmerk
66991
12O0117N

contact
wt@fwo.be
02 550 15 88

datum
22 juni 2016

betreft **Toekenning postdoctoraal onderzoeker 2016-2019**

Geachte heer,

De raad van bestuur heeft u vandaag aangesteld als postdoctoraal onderzoeker van het FWO met ingang van 1 oktober 2016 tot en met 30 september 2019. Het reglement vindt u als bijlage.

Graag willen wij u feliciteren met deze aanstelling.

Voor de verdere administratieve opvolging zal het FWO u nog vóór 15 juli 2016 een arbeids-overeenkomst bezorgen.

Onderzoekers dienen, voor wat betreft de ethische problemen inzake de betrokkenheid van proefpersonen en/of proefdieren, het advies van de lokale ethische commissie in te winnen.

Voor bijkomende inlichtingen kan u via bovenvermeld e-mailadres steeds terecht bij de dossierbeheerder van uw wetenschapsgebied. Vóór 1 oktober 2016 zal u nog een gedetailleerde feedbackbrief ontvangen.

Tenslotte vragen wij u met aandrang om op uw publicaties steeds uw titel "postdoctoraal onderzoeker van het FWO" te vermelden.

Wij wensen u veel succes toe in uw verdere onderzoeksloopbaan.

Met vriendelijke groeten,

Waarnemend secretariaat-generaal,

Danny Huysmans
Directeur Intern Beheer

Hans Willems
Directeur Steun aan Onderzoekers

bijlagen: 1

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)

VASP benchmark on BrENIAC

Kurt Lejaeghere – Arthur De Vos – Sam De Waele

1. Background

BrENIAC contains 580 nodes with 28 cores each, which are of the Broadwell E5-2680v4 type. Each node has 128 or 256 GB RAM and consists of 2 NUMA regions of 14 cores. The network is connected through an Infiniband EDR 2:1 connection.

To benchmark the performance of VASP (module VASP/5.4.1-intel-2016a) on BrENIAC, three very different test systems were considered:

- A doubled $Fe_{16}N_2$ unit cell with one N atom removed
  (35 atoms, 224 bands, 196 irreducible k-points, vasp_std)
  designated by tag METAL

- a Ge semiconductor surface with Pt atoms adsorbed
  (100 atoms, 336 bands, 8 irreducible k-points, vasp_std)
  designated by tag SEMI

- the metal organic framework UiO-66 with two missing linker defects
  (420 atoms, 1120 bands, 1 irreducible k-point, vasp_gam)
  designated by tag PORE

2. Optimal parallelization on 1 node

VASP has the possibility to parallelize over k-points and, for a given k-point, over electronic bands. In general, parallelization over k-points is more efficient, since it requires almost no communication between subprocesses. However, it also substantially increases the memory requirements, since the calculation of the wavefunction at 1 k-point is based on knowledge of all energy levels at that k-point. The memory needed therefore increases when more k-points are computed simultaneously (KPAR). Analogously, parallelization within 1 band occurs by grouping blocks of plane waves in diagonalization routines and allows spreading the memory even thinner. It is more favourable for the memory requirements to devote more cores to a single electronic band (NCORE), equivalent with fewer bands per node, but this behaviour is less distinct.

Table I: Walltime of a calculation of METAL, SEMI and PORE on 1 node, depending on the parallelization settings (number of k-points treated simultaneously, KPAR, and number of cores per band, NCORE).

wall time METAL [s]	NCORE = 1	NCORE = 7	NCORE = 14	NCORE 28
KPAR = 1	9863	6772	6402	6924
KPAR = 2	8654	6515	5601	
KPAR = 4	8435	6369		

wall time SEMI [s]	NCORE = 1	NCORE = 7	NCORE = 14	NCORE 28
KPAR = 1	777	770	710	687
KPAR = 2	775	769	685	
KPAR = 4	720	729		

wall time PORE [s]	NCORE = 1	NCORE = 7	NCORE = 14	NCORE 28
KPAR = 1	4900	4091	4059	3809

Table II: Memory usage per core for a calculation of METAL, SEMI and PORE on 1 node, depending on the parallelization settings (number of k-points treated simultaneously, KPAR, and number of cores per band, NCORE).

mem METAL [MB]	NCORE = 1	NCORE = 7	NCORE = 14	NCORE 28
KPAR = 1	1499	933	883	896
KPAR = 2	2002	1486	1446	
KPAR = 4	3091	2601		

mem SEMI [MB]	NCORE = 1	NCORE = 7	NCORE = 14	NCORE 28
KPAR = 1	328	199	193	187
KPAR = 2	424	307	297	
KPAR = 4	643	525		

mem PORE [MB]	NCORE = 1	NCORE = 7	NCORE = 14	NCORE 28
KPAR = 1	736	406	369	352

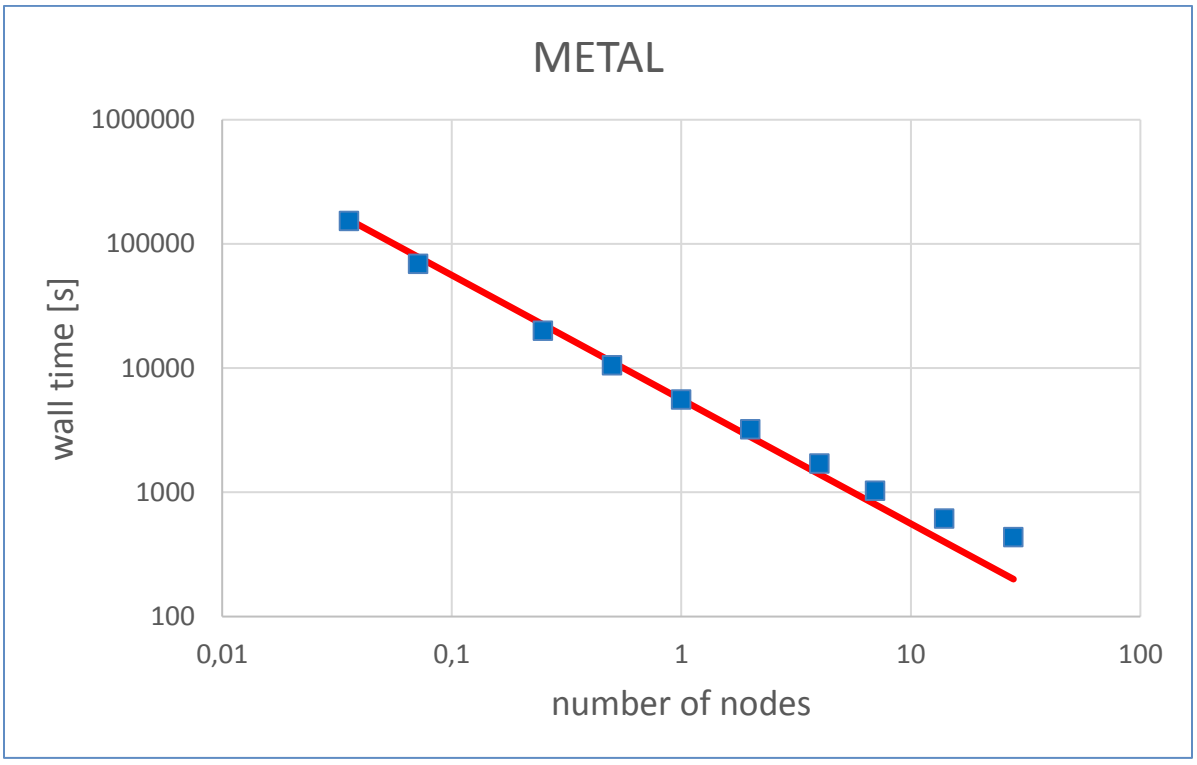
In terms of timing, we see that a higher order of k-point parallelization reduces the required wall time. However, it is not immediately clear which band parallelization is optimal. Many systems benefit from NCORE = 7 or 14, in line with the use of 1 shared memory per band, but for specific systems and number of plane waves, NCORE = 1 may become most favourable (e.g. when increasing the number of plane waves for SEMI). We can only conclude that the *best tradeoff between k-point parallelization and band parallelization needs to be tested for the particular system at hand*. This can be done quite easily, using only a few test calculations (e.g. NCORE = 1, 7, 14 and 28 at KPAR = 1 on 1 node for a representative

system and cutoff energy) and for the optimal configuration taking *KPAR as high as possible*. In addition, the guidelines for memory should be taken into account as well, since *large systems or systems with many k-points (like METAL) may suffer from too high memory requirements*. Finally, the NSIM tag does not matter too much, but NSIM = 1 is strongly discouraged, as it drastically increases the computation time (default is NSIM = 4).

In comparison to Ghent clusters, the (empty) BrENIAC machine performs exceptionally well. For the SEMI system, timings are about two times as good as the best wall times ever achieved on Muk (1378 s in 2013). The same is true in comparison to golett, one of the most recent machines on the UGent HPC (1300 s in 2016). These numbers were scaled to be comparable to the 28 cores per node of BrENIAC. Note, however, that the wall time on golett was measured on the machine in full loading (whereas the BrENIAC machine was almost empty), which has a large impact on the speed of the calculations.

3. Intra- and multinode scaling

Figure 1: Intra- and multinode scaling of the wall time for the METAL system.



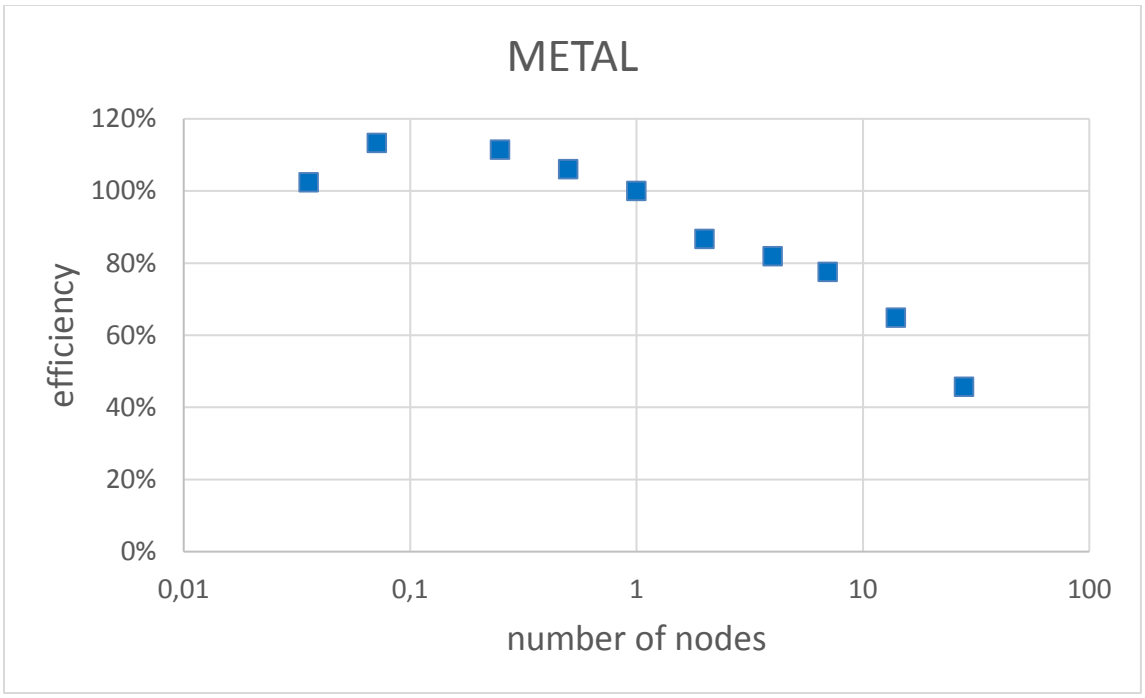
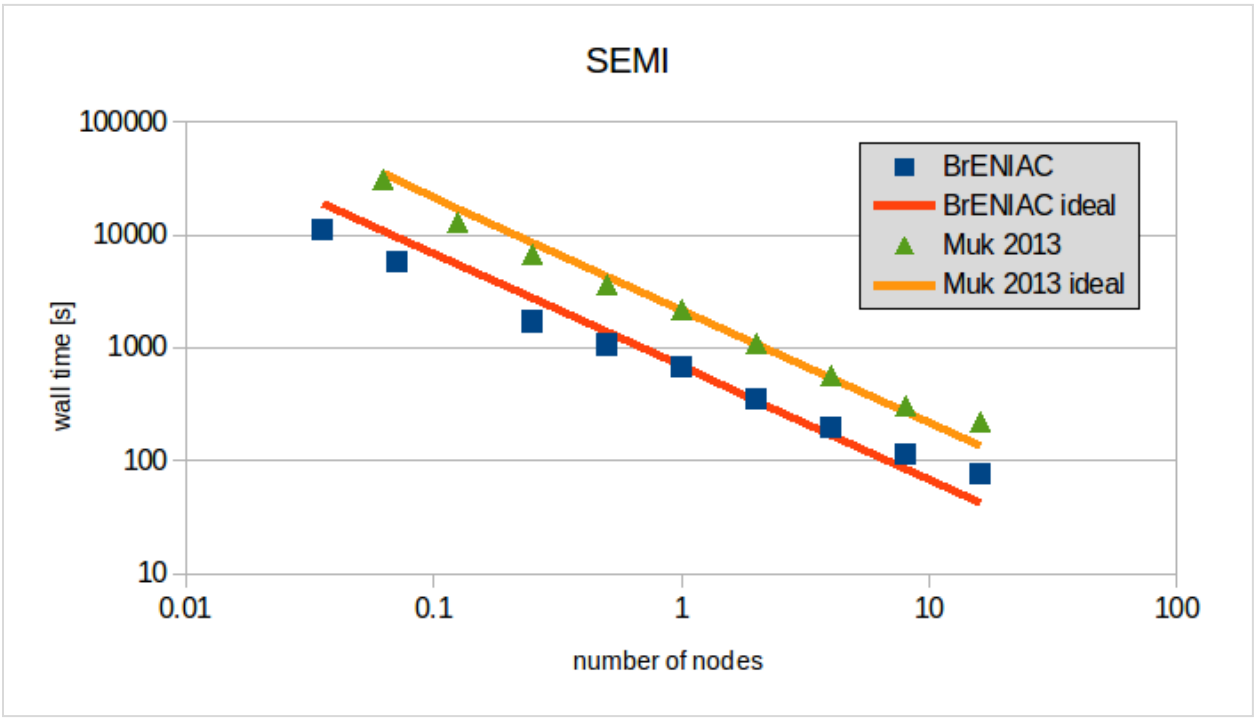


Figure 2: Intra- and multinode scaling of the wall time for the SEMI system (BrENIAC 2016 and Muk 2013). The red and orange lines denote the ideal scaling behaviour.



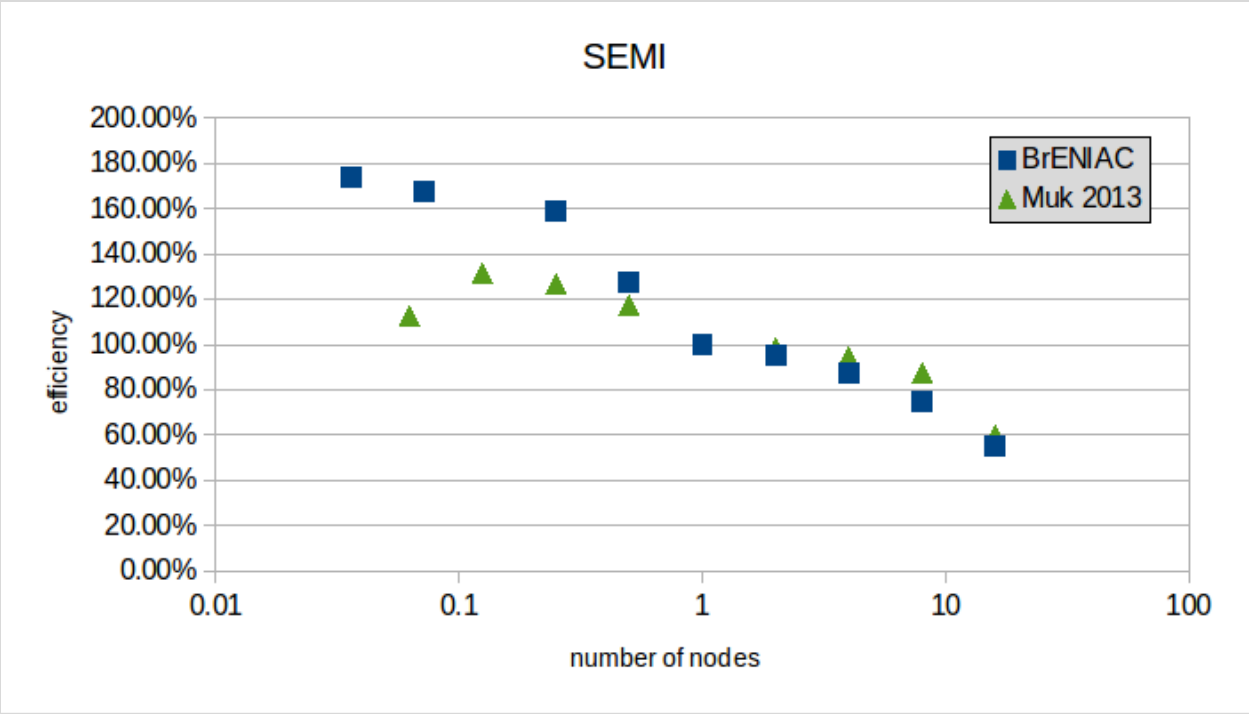
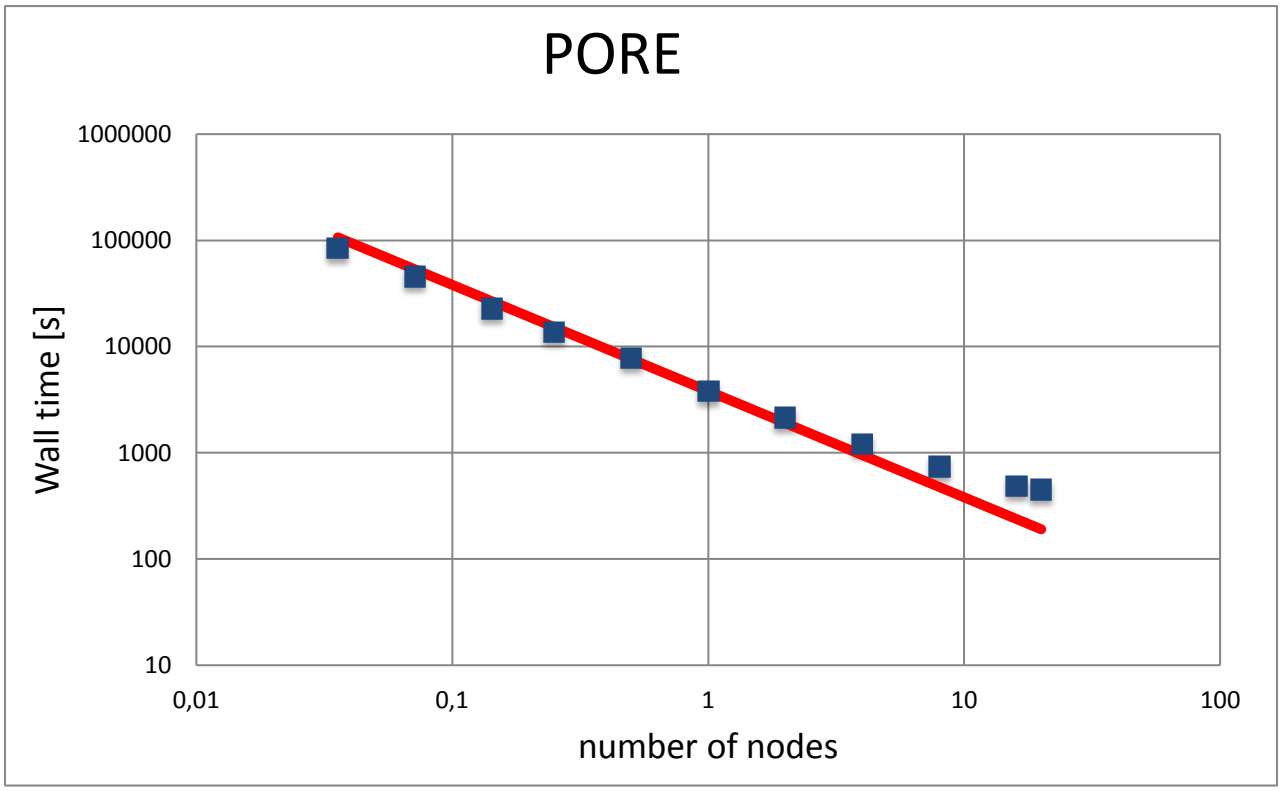
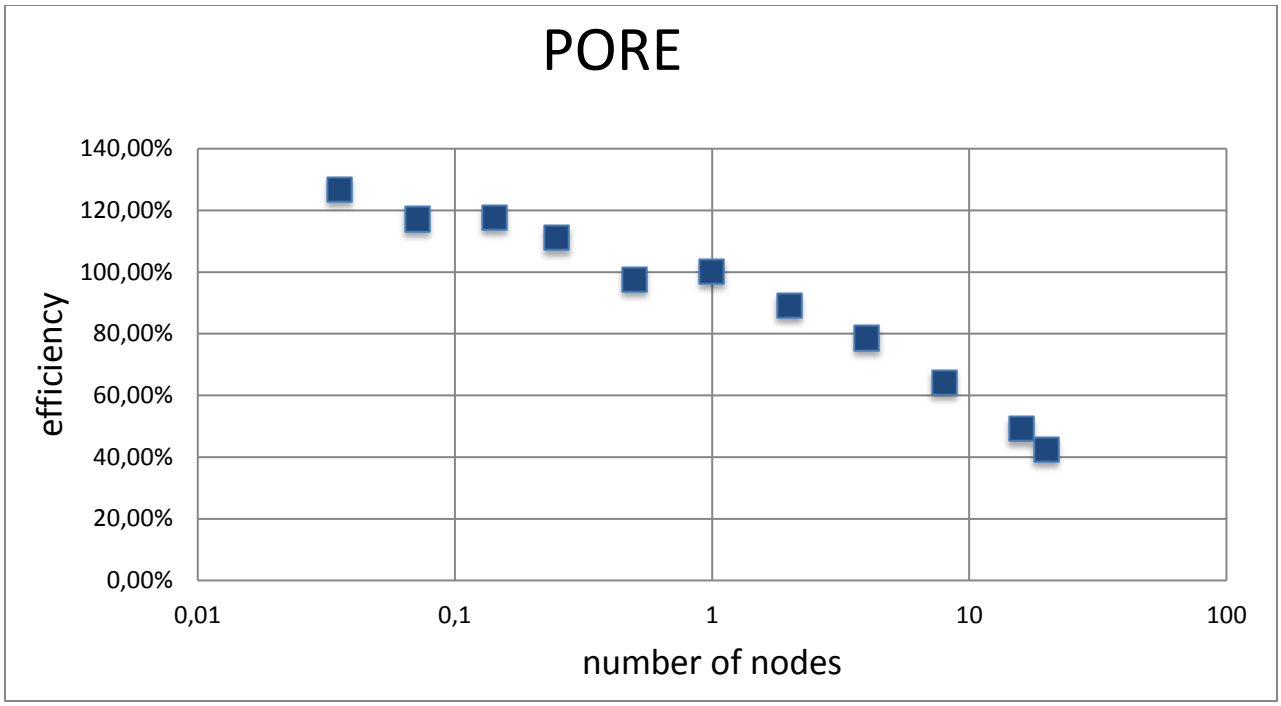


Figure 3: Intra- and multinode scaling of the wall time for the PORE system. The red line denotes the ideal scaling behaviour.





By performing the calculations on a few cores up to multiple nodes, we note that the computational efficiency proceeds in 2 steps. On the one hand, it remains most efficient to perform calculations on 1 or a few cores, and up to the use of an entire node, the efficiency steadily declines. This intranode scaling differs significantly for different systems, however, with poor scaling for SEMI and almost ideal scaling for METAL. *The multinode scaling, however, is quite efficient*, and parallelization over 8 nodes leads to wall times that are still 60-80 % of the efficiency of a single node. Beyond 16 nodes, efficiency drops below 50 %, and calculations are only advisable if they cannot be calculated within 72h on fewer nodes. This behaviour is similar for all tested systems, despite their large diversity, and in line with tests on Muk in 2013 (see Figure 2). We may therefore conclude that it is *not meaningful to perform such scaling tests time and again; only the optimal parallelization settings on 1 node need to be examined when considering a new system*.

As a final note concerning the parallelization settings in multinode calculations, we remark that it is best not to parallelize 1 k-point or 1 band over multiple nodes. Using KPAR equal to the number of nodes (or higher) decreases the computational load significantly, because k-point parallelization requires little communication. For the SEMI system on 2 nodes, for example, a k-point-parallelized calculation (KPAR = 4, NCORE = 14) takes 359 s, while a band-parallelized calculation (KPAR 1, NCORE = 14) takes 408 s. For the METAL system, the difference is huge: 14 002 s for KPAR = 4 and NCORE = 14, compared to 43 229 s for KPAR = 1 and NCORE = 14.