Enclosure 1. Tier-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:

Studying the mechanical and thermal stability of UiO-66 with a full quantum mechanical description

Name and first name of the applicant:

DEMUYNCK Ruben

Institution:

Ghent University

Research group / department:

EA17 / Center for Molecular Modeling

Title / position:

Ir. / PhD Fellow

e-mail address:

Ruben.Demuynck@ugent.be

Total computing time that is needed, in node days:

4320

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

270

The total number of pages in this application should not exceed 17, excluding possible appendices (confirmation letter of financing institution, software license,…) which may be taken into account by the Tier-1 Allocation Board.

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

Combined experimental and computational study of electronically modulated metal-organic framework (MOF) catalysts, FWO project 3G048612 between COK (prof. Dirk De Vos) and CMM (prof. Veronique Van Speybroeck)

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):

Metal-organic frameworks (MOFs) are crystalline, nanoporous materials consisting of inorganic metal nodes connected through organic linkers. Their porous nature endows these materials with an enormous potential in a wide range of applications including sensing, gas separation/adsorption, and catalysis. In this research project, we will investigate one of the most promising MOFs: UiO-66. Compared to other MOFs, UiO-66 shows an exceptionally high thermal stability and retains its crystal structure under high pressures as well as in relatively harsh acidic environments, despite the presence of missing linkers – so-called linker defects. As a result, UiO-66 forms an attractive material for a broad range of applications, including gas adsorption and heterogeneous catalysis. Moreover, its pore size can be tuned by intentionally creating linker defects, incorporating larger linkers, or functionalizing these linkers, paving the path for tailormade materials. In this project, we aim to study the influence of pressure on the formation and stability of linker defects, requiring a full quantum mechanical treatment of the underlying free energy surface.

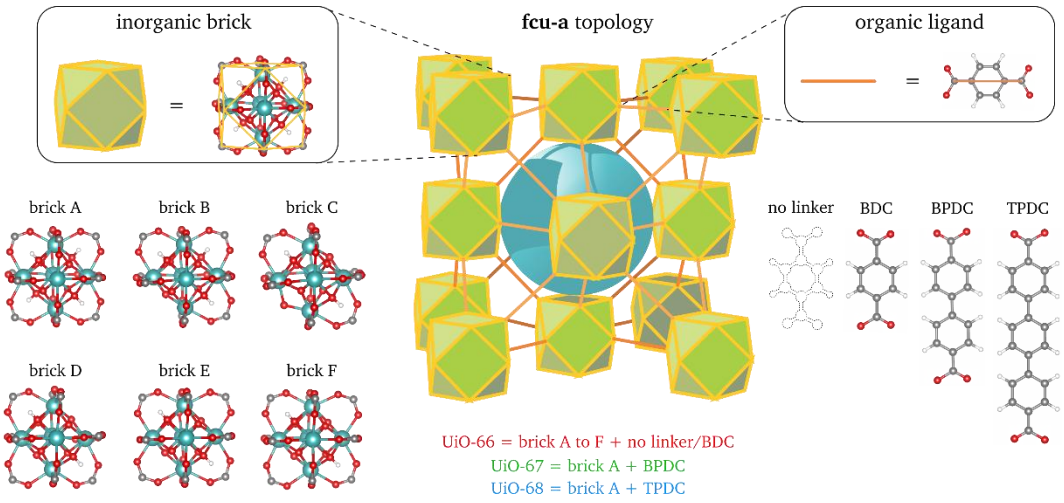


Fig 1.: UiO-66 topology, indicating the physically different defect structures.

Several computational investigations have already been performed on this family of materials, mainly trying to elucidate on the relative stability of the different defect types. However, the number of studies which investigate the

materials with a fully quantum mechanical treatment remains scarce. Recently, we pioneered in this respect by constructing free energy profiles for a breathing MOF, using a full quantum mechanical description of the potential energy profile. In this project, this protocol will be extended to study the stability of UiO-66 under the influence of high pressures and elevated temperatures. Moreover, we will go beyond the perfect crystal structure and study the relation between defects and stability, employing our earlier methodology to classify the defects in physically inequivalent groups (DOI: 10.1021/acs.chemmater.6b01956). To that end, we will use molecular dynamics simulations to explore the most stable structure and their modes of instability for a range of UiO-66 defect structures. Thanks to the quantum mechanical description of the free energy surface, we will for the first time be able to take into account possible bond breaking related with high pressures and elevated temperatures, mimicking the breakdown observed experimentally and allowing us to gain microscopic insight in the thermodynamics associated with defect formation in UiO-66.

3. Provide an engaging abstract (10 lines) for scientific communication on the website in layman's terms. Should this application be bound by a confidentiality agreement (see also item 12 of this application form), provide more details about the specific nature of the confidentiality and indicate why an abstract may not be published.

In this proposal, we will study UiO-66, a metal-organic framework (MOF), showing an elevated thermal and mechanical stability, despite the presence of defects. Inspired by our earlier work, in which we investigated the flexibility of a MOF using a full quantum mechanical description of the free energy surface at operating conditions of temperature and pressure , we aim to further expand our knowledge by studying the influence of defects on the stability on UiO-66 under influence of high pressures. A full understanding of this phenomena unveil the thermodynamics behind the collapse of this promising material, allowing to tailor the material to increase its stability and broadening its possible field of application.

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. In case the project has not gone through a scientific approval process attach a letter of approval of your own institute.

FWO project 3G048612 between COK (Prof. Dirk De Vos) and CMM (Prof. Veronique Van Speybroeck)

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

Prof. dr. ir. Veronique Van Speybroeck
(Veronique.VanSpeybroeck@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)

   DEMUYNCK Ruben
   Ghent University
   EA17 / Center for Molecular Modeling
   ir. / PhD fellow
   vsc40923
   Experience with the local CMM clusters, the HPC UGent Tier-2 clusters, and the Tier-1 machines Muk and BrENIAC

   ROGGE Sven
   Ghent University
   EA17 / Center for Molecular Modeling
   ir. / PhD-FWO Fellow
   vsc40686
   Experience with the local CMM clusters, the HPC UGent TIER-2 clusters, and the Tier1-machine BrENIAC

   VAN SPEYBROECK Veronique
   Ghent University
   EA17 / Center for Molecular Modeling
   prof. dr. ir. / full professor
   vsc40021
   Experience with the local CMM clusters, the HPC UGent Tier-2 clusters, and the Tier-1 machines Muk and BrENIAC

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).

The strategy of studying the stability of MOFs consist in performing a large number of molecular dynamics (MD) simulations. More in particular, we have shown that MD simulations in two ensembles, controlling respectively the cell volume or the pressure (see dx.doi.org/10.1021/acs.jctc.5b00748), yield a complete picture of the mechanical stability of these materials. For both simulation types, a grid is constructed (a volume and pressure grid,

respectively); each grid point requiring one MD simulation that can be run in parallel. From force field simulations on UiO-66, performed at the TIER-2 level, we established 20 grid points as the minimal required number of points. Moreover, to collect relevant statistics, each MD simulation is run for 50 ps (10000 MD steps). Each MD step consists of running one ab initio single-point calculation, the pre- and post-processing of which takes only a limited amount of time. Taking into account the size of the UiO-66 unit cell (228 atoms or 456 atoms for the defect-free primitive and conventional unit cell), such an ab initio single-point calculation becomes quite demanding. Once these two profiles have been simulated for the defect-free UiO-66, this analysis has to be repeated for the eight inequivalent groups of defected structures.

Since the time and length scales exceed the prevailing time and length scales of calculations performed on the Tier-2 clusters and since a large number of these lengthy simulations need to be performed, many multi-node calculations with a long walltime are required, for which the Tier-1 infrastructure is best suited.

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;)
   - memory (maximum 128 GiB/node; 256 GiB/node is available upon motivated request)
   - 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 1 on the next page). Provide additional descriptions of the computing tasks and comments as needed and clearly relate the described tasks to the tasks in the table. 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.

   (typically up to 2 A4 Arial 12).

The aim of this project is to study influence of defects on the stability of UiO-66, which will be investigated based on nine materials: the defect-free UiO-66 and eight unique defected structures. To that end, we subdivide our project in three tasks. Each task consists of running a large number of lengthy molecular simulations in parallel. Due to the large systems under study the total number of simulations steps is rather limited, hence, the corresponding data for each MD run remains rather limited (approx. 0.25 – 0.5 GiB). Making the total data amount easy transferable to the VO data partition at Tier-2. Hence, we aim to run the tasks specified below in a chronological sequence during the project time and after completion of each task copy back relevant data to the TIER-2 infrastructure.

1. Generate input structure and study their relationship:

Before investigating the stability of the defected UiO-66 structures, we should create valid input structures. From our force-field simulations at the TIER-2 level, we have a first idea of the equilibrated input structures at different volume and pressure points. Nevertheless, employing a full quantum mechanical description for the potential energy surface requires to further optimize the atomic positions of the input structures. Such optimization are computationally heavy and thus require two nodes (with default memory) to run for approximately one day. Although being heavy calculations, this type of simulations can be completed within a limited number of steps, hence generating only a limited amount of data (approximately 0.25 GiB disk space for each of the nine structures and each of the 40 grid points). The total number of jobs amounts to 360 (9x40) for these nine UiO-66 structures (20 pressure and 20 volume points, see below).

2. Molecular dynamics simulations at constant volume
Using the input structures generated in the first task, molecular dynamics simulations in the ($NV\boldsymbol{\sigma}_a = \mathbf{0}T$) ensemble will be performed, controlling the volume and temperature. For each structure, such a series of simulations enables us to identify the most stable volume point for that structure at elevated temperature. Moreover, the mechanical stability of the structure can be assessed based on derived properties such as the bulk modulus under operating conditions. Performing these simulations for each of the eight defect structure in addition to the defect-free structure allows us to study the influence of the defect on the absolute stability. The full quantum mechanical molecular dynamics simulations for such large systems are computationally heavy. Hence, for each MD simulation, two nodes need to be run for 5 days with default memory (using restart files to go beyond the maximum walltime of three days), which needs to be multiplied by the number of volume points (20) and structures (9). Since the total number of MD steps is rather limited, the occupied disk space occupied remains limited during the whole series of simulations (approx 0.5 GiB / grid point).

3. Molecular dynamics simulations at constant pressure
Using the input structures generated in the first task, molecular dynamics simulations in the ($NP\boldsymbol{\sigma}_a = \mathbf{0}T$) ensemble will be performed, controlling the pressure and temperature. These simulations will allow us to identify the weakest links in each of the structures, by determining the first point to rupture. Moreover, the elastic tensor can be derived from the cell shape fluctuations during the simulation. Using this elastic tensor, the unique set of three Born stability criteria, which completely determine the stability of the cell, can be derived, allowing us to determine the weakest collective motion responsible for the instability of the material. Performing these simulations for the defect-free and each of the eight defect structure allows us to study the influence of the number and distribution of defects on the instability. The

full quantum mechanical molecular dynamics simulations for such large systems are computationally heavy. Hence, for each MD simulation, two nodes need to be run for 5 days with default memory (using restart files to go beyond the maximum walltime of three days). Multiplying with the number of pressure points (20) and structures (9). Since the total number of MD steps is again rather limited, the occupied disk space occupied remains limited during the whole series of simulations (approx 0.5 GiB / grid point).

Table 1

	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 task	# CPU cores per task	Memory usage (GiB) / node per task	OpenMP / MPI / hybrid	Tier-2 DATA/HOME volume (GiB) + number of files	Tier-1 SCRATCH volume (GiB) + number of files
Generating input structures	9 x 40	1	2	720	56	2	Hybrid OpenMP/MPI	90 GiB + 1440 files	0.25 GiB x 360 + 1440 files
Molecular dynamics simulations at constant volume	9 x 20	5	2	1800	56	2	Hybrid OpenMP/MPI	90 GiB + 720 files	0.5 GiB x 180 + 720 files
Molecular dynamics simulations at constant pressure	9 x 20	5	2	1800	56	2	Hybrid OpenMP/MPI	90 GiB + 720 files	0.5 GiB x 180 + 720 files
TOTAL				**4320**				**270 GiB**	**270 GiB**

.

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 and, if this is not the case, compilation and installation instructions (possibly with reference to existing Tier-2 installation)

   Provide the results of scaling tests that were conducted with this software, preferably on the current VSC 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 run on the current VSC Tier-1, provide the name, architecture, #cores, memory, etc. of the machine that was used to obtain these results and how you think this compares to the current VSC Tier-1. If a different system/problem size is used provide some guidance how it relates to the problem size in the application.
   Provide both a table and scaling plot such as table 2 and plot 1 below (typically up to 3 A4 in Arial 12).

VASP

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

ASE

http://wiki.fysik.dtu.dk/ase/
GNU Lesser General Public License v2.1
Available on BrENIAC

YAFF

http://github.com/molmod/yaff
GNU General Public License
Available on BrENIAC

The main computational cost of this project are the VASP calculations. YAFF and ASE solely serve to automate the VASP calculations in an intelligent fashion. Hence, the overhead cost of this software is negligible and we limit ourselves to providing the scaling properties of VASP.

We did not perform new multinode scaling tests for the current proposal, since our extended benchmark studies (see attachment and Table 2) show a close similarity in scaling behavior for all systems despite their large differences, i.e. both for large and smaller systems, and both for metals and nonmetals. The multinode scaling test (Table 2) for the current proposal are based on the scaling test of UiO-66, another material in the MOF class of materials. This test shows that for 2 or 4 nodes, the intended parallelization, scaling is very good.

The results of some VASP scaling tests, performed on BrENIAC, are available in Table 3 and Plot 1. Tests on 1 node (Table 3) were performed with MIL-53(Al), employing similar settings as will be done during the project. We tested the optimal combination of parallelization over k-points or parallelization over electronic bands (Table 3), as our extended benchmark (see attachment) showed the optimal parallelization settings to be highly system-dependent. In line with previous benchmarks, which showed a high k-point parallelization to be favorable, we will perform all calculations with NCORE = 7 and KPAR as high as possible.

Table 2

# nodes	# cores	absolute timing (s)	speedup	efficiency
1	28	3810	1.00	1
2	56	2061	1.85	0.924
4	112	1176	3.24	0.810
8	196	720	5.29	0.661

16	448	517	7.36	0.460
20	560	478	7.97	0.399
35	980	432	8.82	0.252

Plot 1

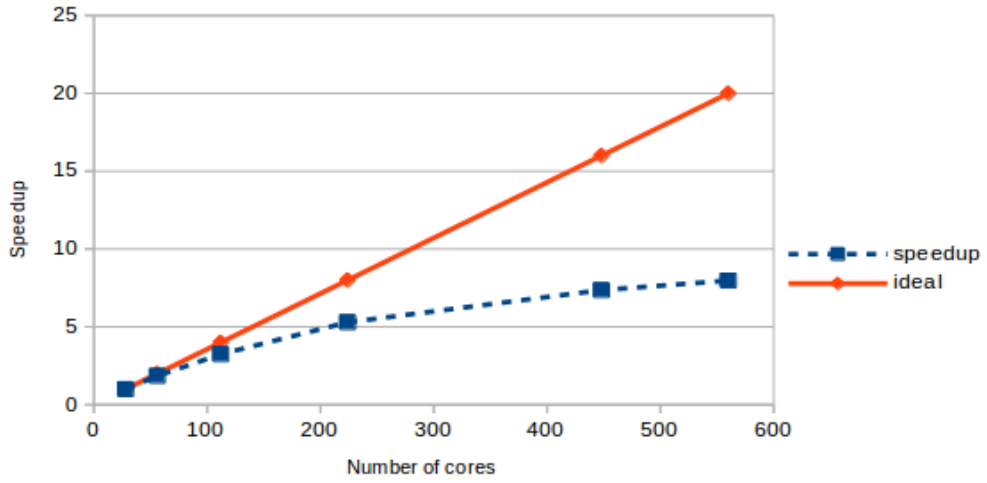


Table 3

absolute timing on 1 node with KPAR = 1 [s]			
NCORE = 1	NCORE = 7	NCORE = 14	NCORE = 28
59.980	46.482	46.085	49.954
absolute timing on 1 node with KPAR = 4 [s]			
56.843	45.510	/	/

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, …)? Provide a data management plan (transfer of files to/from Tier1).

As outlined above, we aim to run our three main tasks in a chronological sense, distributing the time load over the entire time slot from June 2017 until December 2017. The large system under study limits the amount of MD steps performed and hence also limits the total amount of data generated. Thanks to the chronologic work sequence, ordered in three main tasks, our data management plan consists of copying back the data to the TIER-2 infrastructure after completion of each task, limiting the total amount of stored data on TIER-1 to a maximum of 90 GiB.

11. List the granted computing time allocations to the promoter(s) of this research project, on the Flemish Tier-1 systems, 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 systems or on other Tier-1 or Tier-0 supercomputers. DOI links are sufficient.

Ruben Demuynck

[1] Full quantum mechanical study of the influence of functionalization and temperature on phase transformations in metal-organic frameworks (2450 node days)
[2] Construction of ab initio free energy profile for MIL-53-type materials (4200 node days)

Sven Rogge

[1] Full quantum mechanical study of the influence of functionalization and temperature on phase transformations in metal-organic frameworks (2450 node days)

A first manuscript based resulting from the two aforementioned projects has been finalized and will be submitted in the very near future.

Veronique Van Speybroeck

[1] Full quantum mechanical study of the influence of functionalization and temperature on phase transformations in metal-organic frameworks (2450 node days)
[2] Construction of ab initio free energy profile for MIL-53-type materials (4200 node days)
[3] Assessing the accuracy of hybrid functionals for the relative stability of a flexible MOF (3420 node days)
[4] Benchmark study of ab initio molecular dynamics simulations for the methylation of HMB (4812 node days)
[5] High-throughput screening of an unknown quaternary crystal space (4056 node days)
[6] Computational exploration of the free energy profile of guest-free M(bdp) (M=Co,Fe) (bdp2- = 1,4-benzenedipyrazolate) (2940 node days)
[7] Ab initio umbrella sampling simulations of dehydration of UiO-66 (3972 node days)
[8] Assessing the accuracy of an efficient meta-GGA functional for property predictions of elemental solids (640 node days)
[9] Investigating active sites in hydroxylated and dehydroxylated UiO-66 for catalysis of Oppenauer-type oxidation (2110 node days)
[10] Shape tuning of CdSe nanostructures by ab initio determination of the anisotropic growth mechanism (4752 node days)
[11] Dynamical kinetic study of zeolite-catalyzed reactions. (4371 node days)
[12] Structural transformations during dehydroxylation reactions of UiO-66 type metal-organic frameworks; an extension with normal mode analysis. (2726 node days)
[13] Structural transformations during dehydroxylation reactions of UiO-66 type metal-organic frameworks. (4720 node days)
[14] Modeling aldol condensations in metal-organic frameworks with hybrid functional calculations. (2304 node days)
[15] Unraveling reaction pathways on UiO-66 type systems with metadynamics. (4432 node days)
[16] The electronic and magnetic structure of breathing metal-organic frameworks. (4725 node days)
[17] Molecular dynamics study of pentene in H-ZSM-5: towards a better estimate of adsorption enthalpies. (1824 node days)
[18] Exploring the kinetics and selectivity of butene cracking using molecular dynamics simulations. (4864 node days)

[19] Characterizing adsorption properties of C4-C6 alkenes on H-ZSM-5 using molecular dynamics simulations. (4260 node days)
[20] Unraveling dehydroxylation pathways on UiO-66 type systems with metadynamics. (3304 node days)
[21] Dynamical first principle benchmark studies on alkene methylation in H-ZSM-5. (1400 node days)
[22] Investigating the phases of MIL-53-type materials. (4644 node days)
[23] Ab initio molecular dynamics study on the role of water in the reaction mechanism during methanol conversion in H-SAPO-34. (4880 node days)
[24] Defect engineering in UiO-66: How linker defects affect the electronic structure. (4116 node days)
[25] DFT study of reaction paths in zeolite-catalyzed 2-hexene cracking. (4536 days)
[26] Dynamical first principle modelling of zeolite dealumination in H-SSZ-13. (3624 node days)

Articles:
http://dx.doi.org/10.1039/C4RA16800C
http://dx.doi.org/10.1039/C4CE01672F
http://dx.doi.org/10.1002/chem.201500473
http://dx.doi.org/10.1016/j.jcat.2015.01.013
http://dx.doi.org/10.1002/cctc.201402146
http://dx.doi.org/10.1021/cs400706e
http://dx.doi.org/10.1039/c4mh00127c
http://dx.doi.org/10.1039/C3CP54132K
http://dx.doi.org/10.1021/acs.jpcc.5b06809
http://dx.doi.org/10.1016/j.jcat.2015.08.015

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

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: hpc@vub.ac.be

For the other institutions: caroline.volckaert@FWO.be

Fonds Wetenschappelijk Onderzoek
Research Foundation – Flanders

Prof. Dirk De Vos
Katholieke Universiteit Leuven
Oppervlaktechemie en Katalyse
Kasteelpark Arenberg 23
3001 LEUVEN

uw kenmerk	ons kenmerk	contact	datum
	05_05 G.0486.12	bio@fwo.be cult@fwo.be wt@fwo.be gm@fwo.be med@fwo.be interdisciplinair@fwo.be	16 november 2011

Betreft: **Toekenning project G.0486.12**

Geachte professor,

De Raad van Bestuur van het FWO heeft uw bovenvermelde aanvraag voor een onderzoeksproject goedgekeurd. De bijgevoegde overeenkomst werd opgesteld volgens het reglement van de Onderzoeksprojecten van het FWO.

Graag wil ik u hierbij feliciteren met de toekenning van het door u aangevraagde onderzoeksproject.

Mag ik u vragen alle exemplaren te ondertekenen, één kopie voor u te bewaren en mij het origineel samen met de overige kopie(ën) per kerende post terug te sturen.

De toegekende toelage is beschikbaar voor zover de financierende overheden hiervoor de nodige middelen vrijmaken. Onderzoek waarvoor medisch ethisch advies noodzakelijk is, kan pas starten als de onthaalinstelling hierover positief advies uitbrengt.

Slechts uitzonderlijk kan, met akkoord van het FWO, de oorspronkelijke bestemming van dit krediet worden gewijzigd.

Tenslotte vraag ik u met aandrang om op uw publicaties steeds te vermelden dat ze tot stand kwamen met de financiële steun van het FWO. Vacatures gekoppeld aan onderzoeksprojecten kunnen bekend worden gemaakt via de FWO-website. U kan het vacaturebericht aan het FWO bezorgen via communicatie@fwo.be. Aanvullend kan het nuttig zijn om dergelijke berichten ook te publiceren via de Euraxess website van de Europese Commissie: http://ec.europa.eu/euraxess/ .

./..

EGMONTSTRAAT 5
1000 BRUSSEL
TEL +32 2 512 91 10
FAX +32 2 512 58 90
E-MAIL: POST@FWO.BE
WWW.FWO.BE

../..

Fonds Wetenschappelijk Onderzoek
Research Foundation – Flanders

Voor bijkomende inlichtingen kan u steeds terecht bij uw dossierbeheerder. De e-mailadressen zijn voor de Biologische Wetenschappen: bio@fwo.be, voor de Cultuurwetenschappen: cult@fwo.be, voor de Gedrags- en Maatschappijwetenschappen: gm@fwo.be, voor Wetenschap en Technologie: wt@fwo.be, voor de Medische Wetenschappen: med@fwo.be en voor het Interdisciplinair onderzoek: interdisciplinair@fwo.be.

Hoogachtend,

dr. ir. Elisabeth Monard
secretaris-generaal

Bijlagen: 2

Kopie: Prof. Veronique Van Speybroeck

EGMONTSTRAAT 5
1000 BRUSSEL
TEL +32 2 512 91 10
FAX +32 2 512 58 90
E-MAIL: POST@FWO.BE
WWW.FWO.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)

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

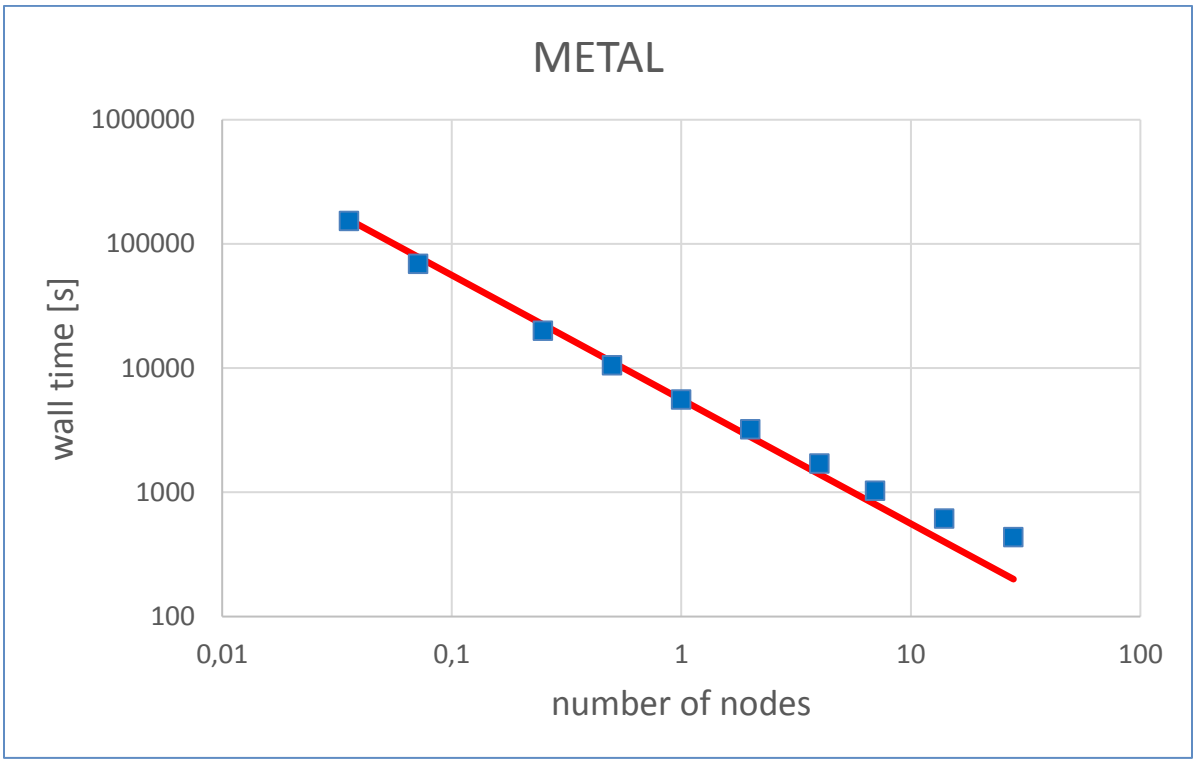
In terms of timing, we see that a higher order of k-point parallelization reduces the required wall time. However, it is not immediately clear which band parallelization is optimal. Many systems benefit from NCORE = 7 or 14, in line with the use of 1 shared memory per band, but for specific systems and 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.



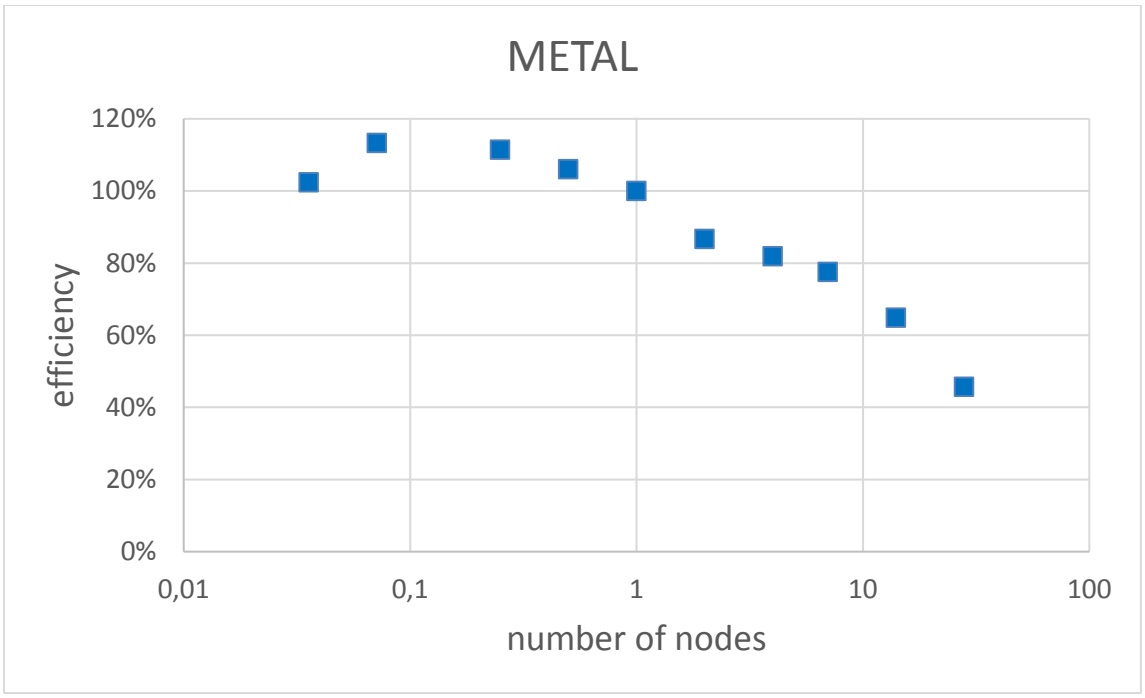
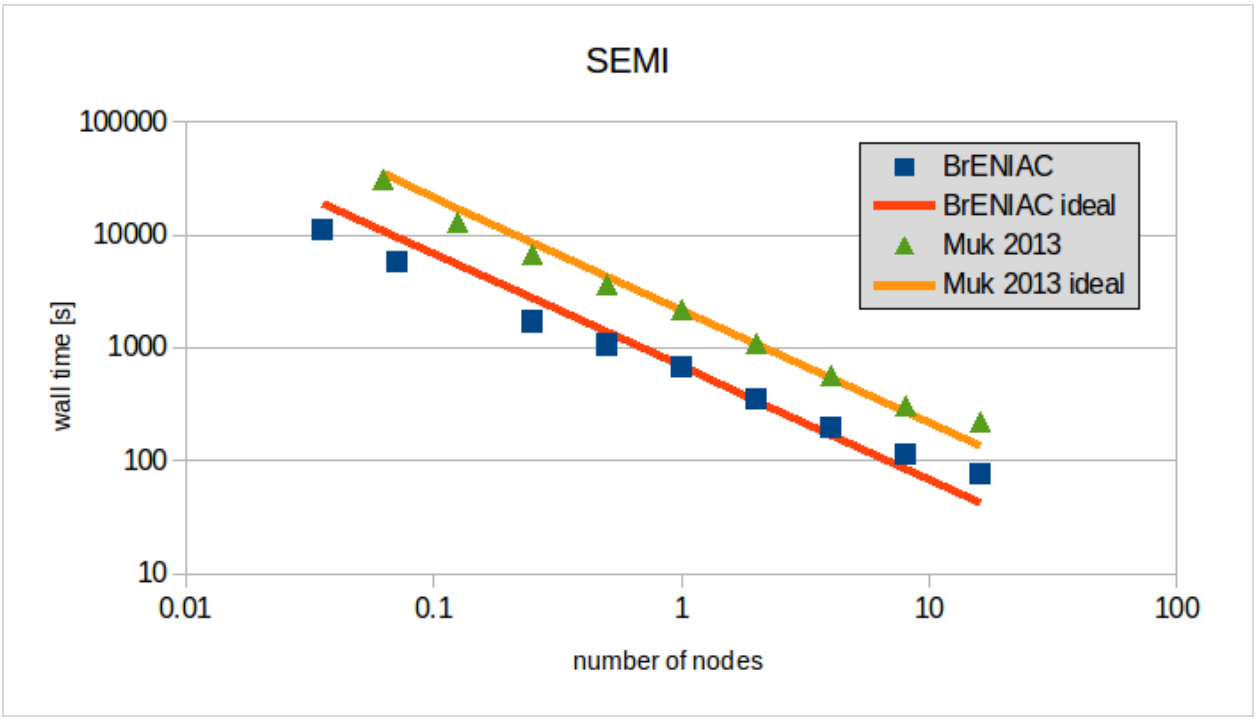


Figure 2: Intra- and multinode scaling of the wall time for the SEMI system (BrENIAC 2016 and Muk 2013). The red and orange lines denote the ideal scaling behaviour.



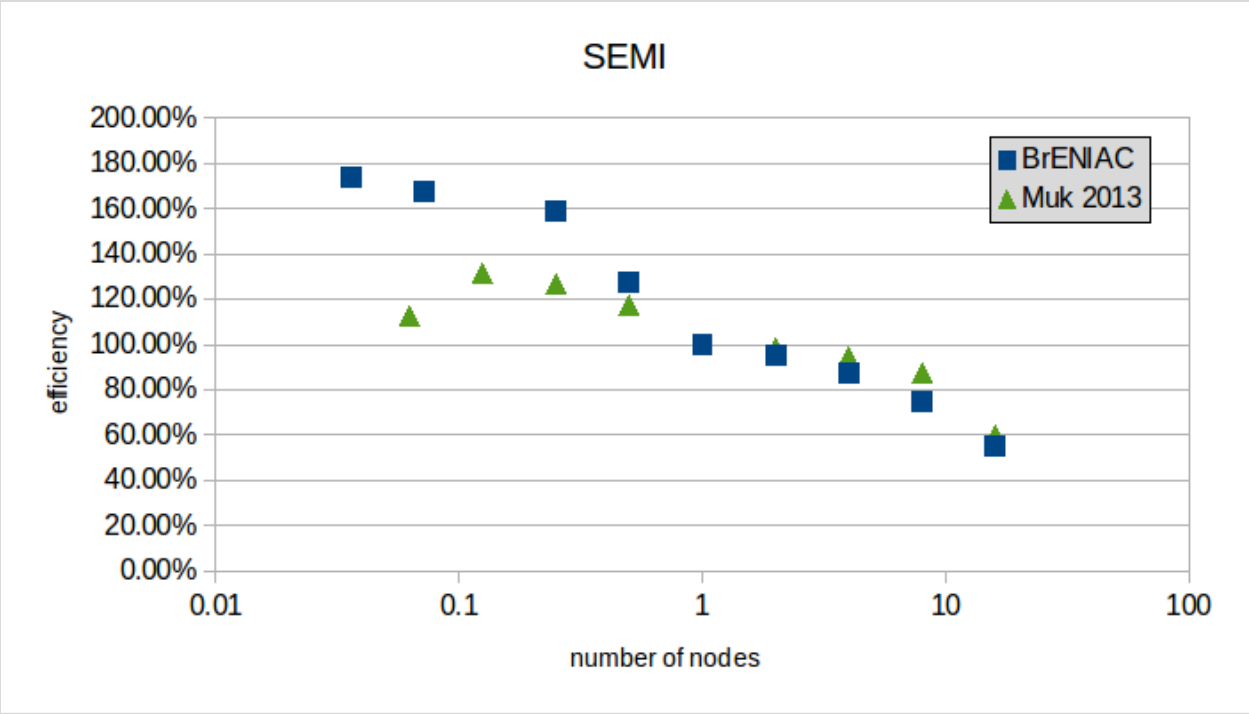
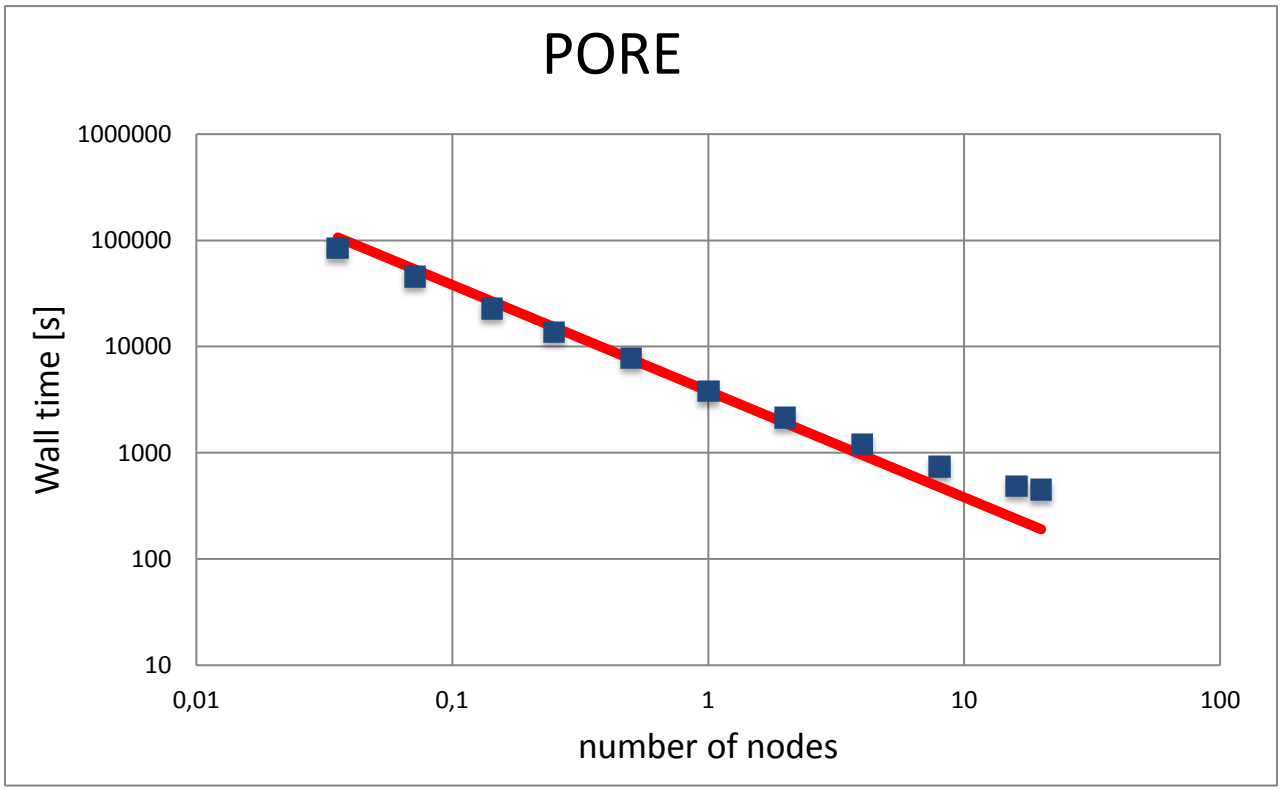
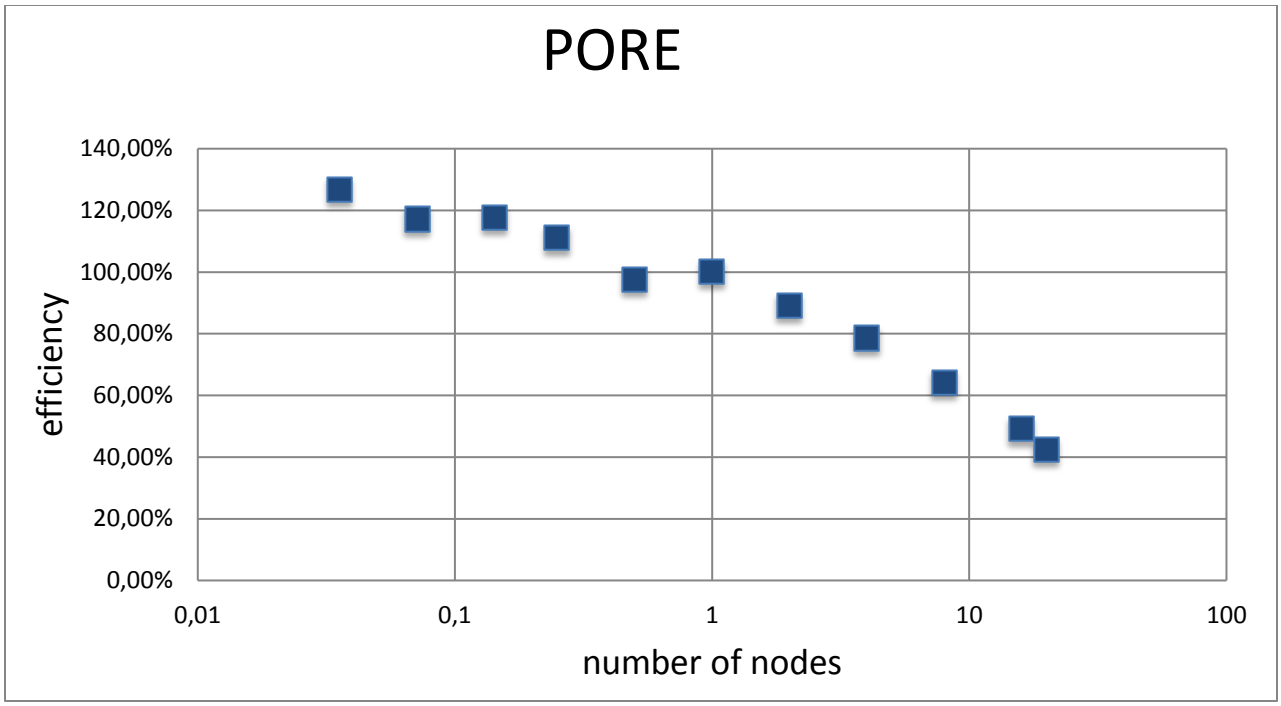


Figure 3: Intra- and multinode scaling of the wall time for the PORE system. The red line denotes the ideal scaling behaviour.





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.