Expressions of Interest for PRACE Programme Access for the MAPPER project

Multiscale modelling & Simulation is quickly turning into an important paradigm in many communities, ranging from the life sciences to the physical & engineering sciences. This is very much driven by the availability of extensive datasets on all relevant spatio-temporal scales, but also by the ever increasing computational capabilities as offered by existing and emerging e-Infrastructures.

One important aspect of many multiscale models, cutting across most communities and application domains, is that they can be composed as a collection of single scale models coupled through the scales. This means that despite the obvious differences in terms of single scales models and multiscale coupling methodologies that exist between scientific domains, we are able to identify generic execution profiles for multiscale simulations. This we call *Multiscale Computing*. Moreover, it turns out that many state-of-the-art multiscale models contain one or more single scale models that have computational needs on the PRACE Tier-1 or even Tier-0 level, while the other single scale kernels are perfectly fit for Tier-2 resources as offered by community clusters, clouds or grids.

These observations, combined with our wish to exploit the full European HPC ecosystem, have led to a European initiative to develop an environment for *Distributed Multiscale Computing* (DMC). This initiative is currently funded through the MAPPER project.[1]

The MAPPER project has a portfolio of seven large scale applications from five domains (Virtual Physiological Human –VPH–, Fusion, Computational Biology, Nanomaterial science, and Hydrology). Each of these applications contains 2 to 5 coupled single scale models, of which typically 1 or 2 require high-end computational resources (PRACE Tier-1 or Tier-0).[2] The VPH applications revolve around cardiovascular disease (coronary and inter-cranial), the Fusion applications aim at simulations of different issues of ITER, the Nanomaterial science application is on simulations of properties of layered nano-structured materials, the Hydrology application is about flow in complex irrigation systems, and Computational Biology is on multiscale optimizations for gene regulatory networks.

Driven by this application portfolio, viewed as a representative subset for multiscale applications, MAPPER aims to deliver an infrastructure that supports DMC on – and across European e-Infrastructures (EGI, PRACE). MAPPER relies as much as possible on existing software and services available via these e-Infrastructures, or as delivered by earlier projects. On December 7, 2010 the MAPPER project sent a letter to both PRACE and EGI with a proposal to explore the idea of DMC and to seek collaboration between all three projects. This resulted in a taskforce between MAPPER, PRACE and EGI, that will demonstrate DMC applications executing on EGI and PRACE resources at the same time using new capabilities for multi-site advance reservation and co-allocation. Additionally, various new application tools for DMC will be deployed and supported. This will be an important technological milestone that will be reached in November 2011.

For the next phase MAPPER needs to move to stable solutions for DMC, providing production level services, so that the full application portfolio is able to perform production runs in DMC mode, with the goal to produce high impact scientific results as well as operational experience with - and statistics on DMC. To do so, MAPPER needs to secure sufficient access to e-Infrastructures. MAPPER has already gained access to a number of grid resources via support from National Grid Initiatives in Poland, Germany, and UK.[3] However, MAPPER also needs to secure access to PRACE resources. With this expression of interest we apply for program access mainly to Tier-1 systems, and a small

---

[1] MAPPER is funded by the eInfrasructure unit of the European Commission under grant number 261507, see www.mapper-project.eu

[2] For details we refer to the MAPPER website.

[3] MAPPER also aims to get access to Dutch and Nordic grid infrastructures.

allocation on Tier-0 systems for the full MAPPER application portfolio[4]. We believe that program access is required because coordinating the deployment of production DMC simulations using more than one site may require a larger timeframe than the 1-year window for project access proposals. Given the added science potential of distributed multiscale simulations, this larger timeframe is justified, but can only be well accommodated in the form of program access. The MAPPER project will run for another two years, ending in September 2013. We therefore propose to offer the MAPPER project PRACE Program Access for its application portfolio for the remaining duration of MAPPER. The majority of the program access should then be for Tier-1 resources, and only limited-size allocations on Tier-0 resources. We foresee that some applications in MAPPER may exceed the scope of the program access (requiring substantial Tier-0 resources). In those cases individual groups should be allowed to apply for additional Project Access.

To support the MAPPER portfolio in the remaining two years of the project we ask for 20 MCPU Hrs per year, with 17 MCPU Hrs per year on Tier-1 resources.

The community allocation will be managed by a small team of MAPPER participants (Prof. P. Coveney, Dr. D. Coster, Dr. A. Hoekstra). This panel will monitor progress and review outcomes relying on bi-annual reporting as already in place in the MAPPER project.

MAPPER staff will continue to work with PRACE and EGI in the taskforce to further develop and demonstrate DMC, while at the same time producing key science output for its applications. Important requirements for DMC are capabilities for advanced reservation and co-allocation.

MAPPER is very well aware of the non-technological hurdles that the vision of DMC needs to face. These are a.o. related to access policies (e.g. advanced reservations and co-allocation of disparate sets of resources). MAPPER has written a report on this[5] and is active in putting these issues at the right agendas (by being active in e.g. the eIRG, eInfra concertation meetings, etc.). We believe that these issues require much more time and discussions, but should not stop us from our ambition to make the vision of DMC a reality. Program Access for the MAPPER project is hgkey to demonstrate the potential of DMC, and to should help in further developing a vision of a European HPC eco system with seamless access to the full computing pyramid, from Tier-2 to Tier-0 and back, and in case of DMC (or other coupled applications), easily coupling resources on all performance levels in the eco system.

Bibliography

Evans DJW, Lawford PV, Gunn J, et al. The application of multiscale modelling to the process of development and prevention of stenosis in a stented coronary artery. Phil.Trans. R. Soc. A 2008; 366: 3343-3360.

Hoekstra, A.G. and P. Coveney. Towards Distributed Multiscale Computing for the VPH. in VPH2010, Brussels, 2010.

Hoekstra AG, Lorenz E, Falcone J-L, et al. Towards a Complex Automata Framework for Multi-scale Modeling. International Journal for Multiscale Computational Engineering 2007; 5: 491-502

Sloot, P.M.A. ,A.G. Hoekstra, Multi-scale modelling in computational biomedicine. Brief Bioinform, 2010. 11, 142-152.

Mazzeo MD, Coveney PV. HemeLB: A high performance parallel lattice-Boltzmann code for large scale fluid flow in complex geometries. Computer Physics Communications 2008; 178: 894-914.

Greenwell HC, Jones W, Coveney PV, et al. On the application of computer simulation techniques to anionic and cationic clays: A materials chemistry perspective. Journal of Materials Chemistry 2006; 16: 708-723.

Suter JL, Anderson RL, Greenwell HC, et al. Recent advances in large-scale atomistic and coarse-grained molecular dynamics simulation of clay minerals. Journal of Materials Chemistry 2009; 19: 2482-2493.

---

[4] Note that in the final phase of the MAPPER project the aim is to also allow external applications to access and use the MAPPER infrastructure. MAPPER will work with a light weight review before allowing a limited number of external application to use the infrastructure.

[5] MAPPER deliverable D3.1: Report on the policy framework resource providers need to adopt to support the MAPPER Project, available at http://tinyurl.com/6zechlf .