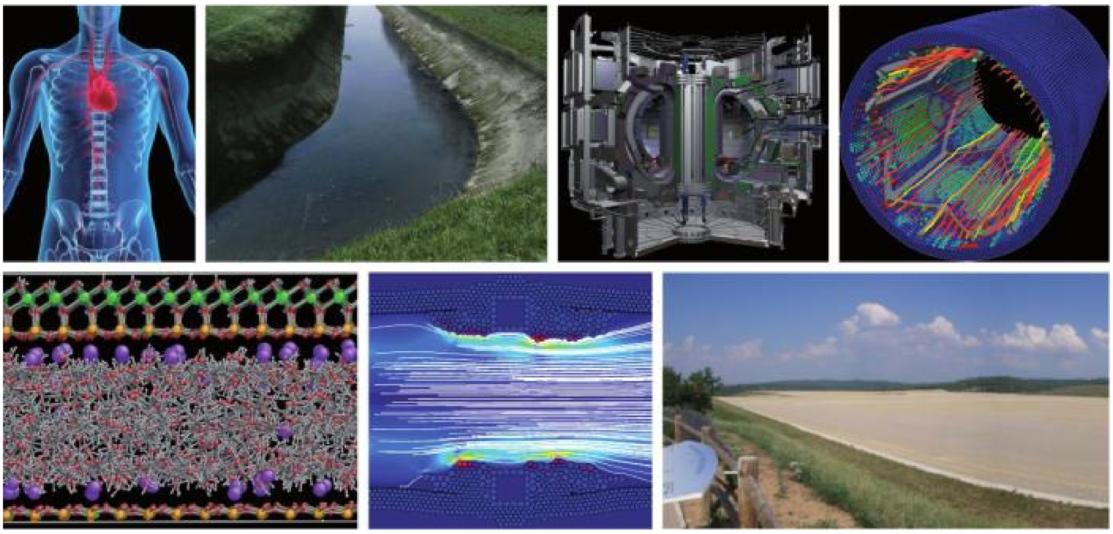
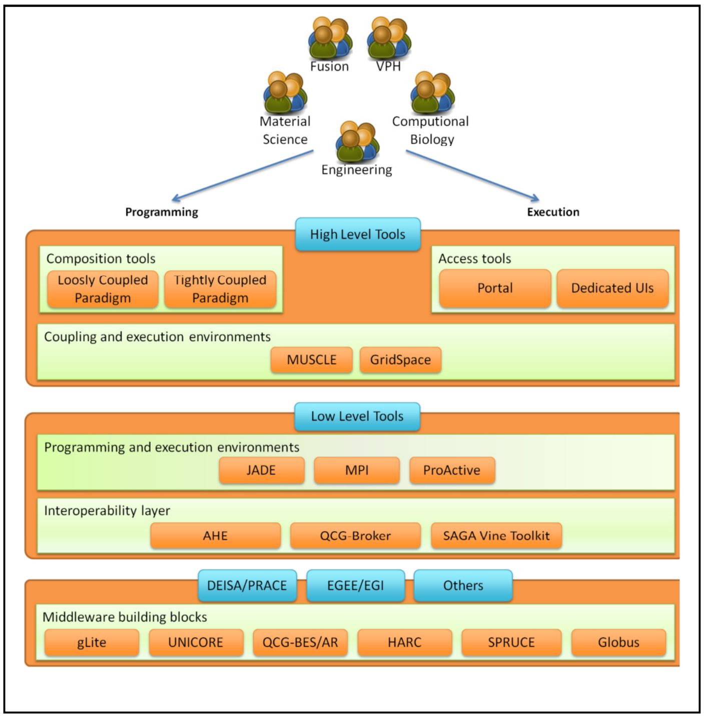
# Status update

*An overview of the vision and update on the results after two years*



www.mapper-project.eu

## PROJECT CONTEXT AND OBJECTIVES



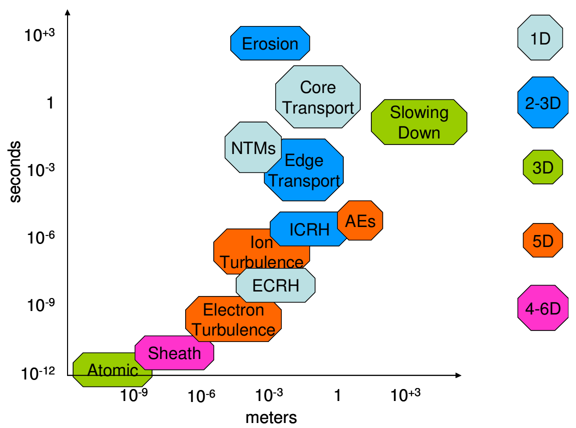

Driven by seven challenging applications from five representative scientific domains (fusion, clinical decision making, systems biology, nano science, hydrological engineering), MAPPER deploys a computational science environment for distributed multiscale computing on and across European e-infrastructures. By taking advantage of existing software and services, as delivered by EU and national projects, MAPPER will result in high quality components for today's e-infrastructures. We develop tools, software and services that permit loosely and tightly coupled multiscale computing in a user friendly and transparent way. We integrate our applications into the MAPPER environment, and demonstrate their enhanced capabilities.

MAPPER integrates heterogeneous infrastructures for programming and execution of multiscale simulations. The MAPPER solutions are developed on top of existing e-infrastructures without the necessity to modify already deployed components.

We distinguish two layers of services constituting the MAPPER environment. Users and applications communicate with services belonging to the interoperability layer, an abstract layer to grid resources managed by different middleware stacks. The interoperability services are responsible for providing concurrent access to resources controlled by different services synchronizing and orchestrating the execution of applications in the grid.

Multiscale loosely and tightly coupled simulations are controlled by a broker and by underlying computing access services developed in the FP6-ICT QosCosGrid project (http://www.qoscosgrid.org/). The broker is integrated with underlying middleware and its scheduling and co-allocation algorithms are tuned for specific needs of multiscale applications. Many of the services that we wish to use have been developed individually and do not necessarily interoperate. We ensure that these services do talk to each other where appropriate. We adopt a twin track approach in our service development activities. The fast track adapts, integrates and deploys a minimal set of infrastructure components to enable coupling of multi-scale applications. The deep track will do so for the higher level services required to realise the full and integrated MAPPER infrastructure, which will enable the coupling and launching of multi-scale component codes. MAPPER services evolve on the basis of a regular cycle of top-down and bottom-up analysis of existing e-infrastructure, MAPPER building blocks as well as new requirements defined by our multiscale user communities.

The application portfolio is adapted to the MAPPER infrastructure. Our approach is that applications are up and running from the start of the project, with existing, easily adaptable and deployable tools in the fast track; the deep track then produces enhancements, which are fed into the user level fast track as and when ready. A number of programming and execution tools, dedicated to distributed multiscale computing, are developed. In the first phase of the project, the applications will have to rely on explicit coding of their multiscale simulations, but gradually programming tools are delivered that assist in this task.




Range of scales and processes involved in the fusion applications.

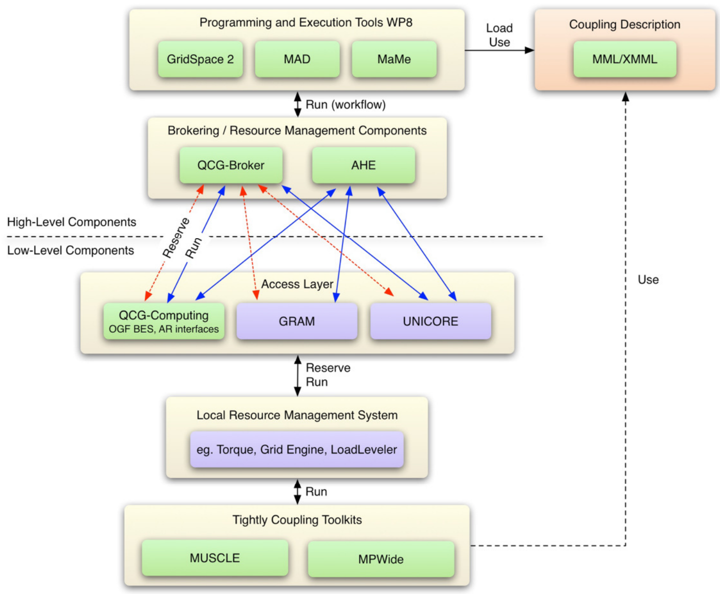
Although MAPPER is driven by seven exemplar applications from five user communities, our solutions are generic and will enable distributed multiscale computing for any multiscale model fitting into our paradigm, and MAPPER therefore opens up to other user communities as well.

MAPPER partners have significant trans-Atlantic grid and HPC experience, and have been involved very actively in TeraGrid and with the US Department of Energy laboratories. We collaborate with the US TeraGrid to integrate infrastructures across the globe.

## MAIN RESULTS AFTER Y2

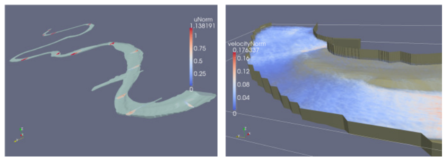
The first four months of the project were dedicated to a complete characterisation of the MAPPER application portfolio. This description includes details on the coupling templates, the computational demands of the sub models and the technical resource requirements for running the application as a distributed multiscale simulation. Moreover, a more generic user needs analysis, requirements analysis, and an assessment of current capabilities of existing e-infrastructures was carried out. This resulted in a major milestone for the project, and the deliverable describing this work has laid the groundwork for the sequel of the project.

During the second half of the first year, a few major technical results were obtained. A first refactoring and adaptation of the applications, to meet the requirements of the MAPPER infrastructure, was performed. A design of the programming and execution tools was delivered, and a first prototype implementation of the tools released. The MAPPER-EGI-PRACE taskforce was established. This taskforce has become key in achieving both technical goals of MAPPER as well as achieving sustainability goals of MAPPER. Key MAPPER services were deployed on all resources of the MAPPER pre-production infrastructure and selected EGI and PRACE resources. Finally, two distributed multiscale applications, deployed and managed on e-Infrastructures in Europe, including adaptation a first vertical integration of key MAPPER services, were demonstrated by the end of the first year.




Mapper vertical Integration

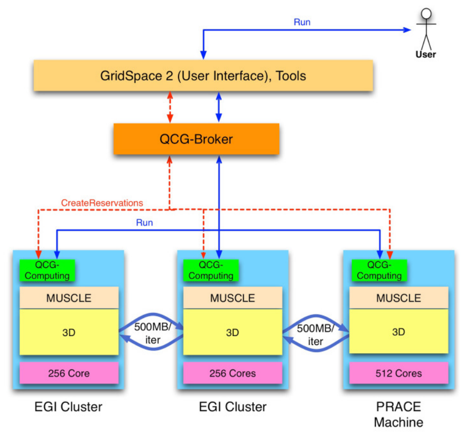
The main goal of the second year was to have all applications operational and running on the MAPPER infrastructure. This has been achieved. All applications run in Distributed Multiscale Computing mode on the MAPPER infrastructure, in a variety of scenarios (running on EGI resources, on a mix between PRACE and EGI, both in loosely coupled and tightly coupled mode.



Simulation of the flow in the Rhone river in Geneva (left). Example of a 3D, free surface simulation of a draining event at the dam in Verbois (right).

The full MAPPER tool chain has been used to achieve this result, and in doing so, providing valuable feedback on the tools themselves. In combination with feedback achieved during the first MAPPER seasonal school, the project has now improved the tools both in functionality and usability.

To facilitate distributed multiscale computing the project has continued its effort to deliver, integrate and deploy dedicated services. This deep track development has resulted in a number of improved and adapted key services, such as MUSCLE 2.0, and a full integration of the 4 main MAPPER services (QosCosGrid, AHE, GridSpace, MUSCLE).



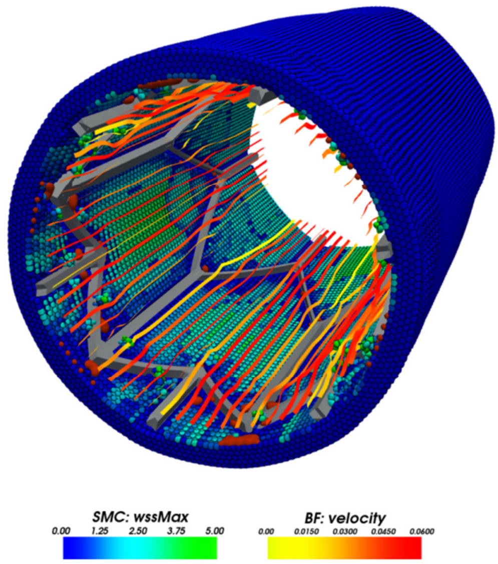

Irrigation Canals Application Scenario

A key result of MAPPER is the continued collaboration between PRACE, EGI and MAPPER in the taskforce. This not only resulted in a number of important results (distributed multiscale computing between PRACE and EGI sites for both loosely coupled and tightly coupled scenarios), and exciting proofs of concept (e.g. advanced reservation on PRACE sites), but most important, a first major result in sustainability. Due to the intense continued collaboration between MAPPER and EGI in the taskforce, a major MAPPER service (the QosCosGrid software stack) has now passed all thresholds and is being integrated in the standard EGI software stacks. The QosCosGrid stack was integrated with EGI management, help-desk, monitoring and accounting infrastructures. All the efforts were formalized by signing a Memorandum of Understanding between EGI.eu and PSNC (acting as the QosCosGrid middleware Technology Provider) in October 2012. This collaboration and integration of MAPPER services has also been described and acknowledged in EGI-Inspire deliverables. WE believe that the taskforce has also resulted in a focussed collaboration between EGI and PRACE and we intend to foster this for the next year of MAPPER. Not only to further work on achieving MAPPER's technical goals, but also as a route towards sustaining MAPPER outcomes.

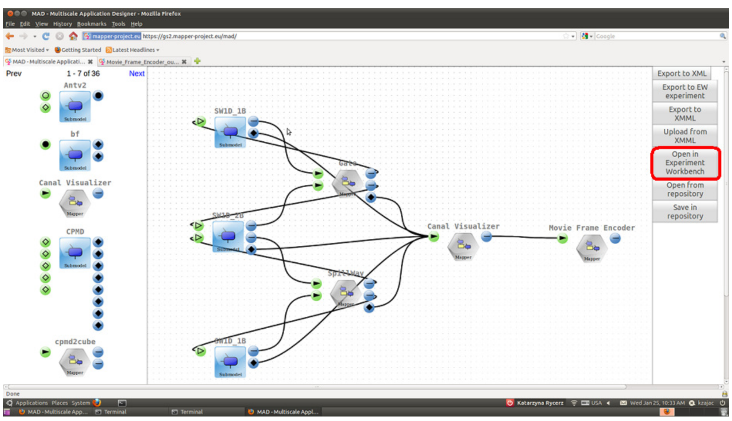
MAPPER's dissemination actions have resulted in increased visibility, with papers in scientific journals (7), and posters (3) and (invited) lectures (8) in scientific events. Moreover, MAPPER had dedicated booths both at generic computing events (such as ISC2012) and domain specific conferences (such as VPH2012). MAPPER organised it's first seasonal school in January 2012 in Londen. The school attracted 25 participants with 36% external delegates, mainly from from related projects such as MMS@HPC, Scalalife, and Shiwa. Due to MAPPER's increased visibility we now observe more and more awareness of the solutions as offered by MAPPER, resulting in requests from thirds parties to start using MAPPER services and infrastructures. This we will capitalize during the 3rd year of the project.

## EXPECTED RESULTS AND IMPACT

The main outcome of the MAPPER project will be a computational science software infrastructure that will drastically lower the barrier to compose, adapt and maintain multiscale simulations, and, most important, to seamlessly deploy and execute such simulations on European e-Infrastructures. We expect to see immediate impact in the scientific communities that we specifically address via our applications portfolio. Moreover, through our anticipated dissemination and outreach actions we will gradually enlarge the use base of MAPPER tools. Moreover, by starting to work on sustainability of MAPPER developments immediately at the start of the project, we expect that MAPPER can really live up to expectations by offering sustained solutions after the lifetime of the project.



A result from an in-stent restenosis simulation run


A screenshot from the MAD tool

The software, methods and infrastructure we develop will be of great use to many computational scientists, whether academic or non-academic, regardless of scientific discipline. Computational science-based research is ubiquitous across academia and most areas of manufacturing and extraction industries, finance, commerce, and government – with innovative, novel or large scale computing resources (including computational grids, clusters or high performance computing (HPC)) becoming increasingly prevalent. With the relentless growth in computer power, ever more complex problems become solvable using advanced computer simulation, so the demand for its methods, together with a well-trained workforce with the skills to exploit such capabilities, will increase over the coming decades.

Mapper booth during VPH2012 in London, September 2012

We recognize that facilitating the type solutions that we propose for distributed multiscale computing is not only of a technical nature. Access policies to infrastructure may prohibit otherwise technically sound solutions. We therefore also dedicate some project effort to identify potential issues, and on the right policy levels intend to start lobbying in order to attain maximal MAPPER impact as well.

MAPPER capitalizes on earlier EU funded project by using, extending, adapting and improving software developed in those projects. The software components developed within the project will be distributed under an Open Source license. MAPPER will train young computer scientists in developing and maintaining high quality complex software, thus contributing to a European wide high quality human capital in software engineering.