Deliverable D3.6

Final Sustainability Plan

Project acronym: MAPPER

Project full title: Multiscale Applications on European e-infrastructures.

Grant agreement no.: 261507

Due-Date:	30.9.2012
Delivery:	30.9.2012
Lead Partner:	LMU
Dissemination Level:	PU
Status:	Final
Approved:	Q Board, Project Steering Group
Version:	1.0

DOCUMENT INFO

06.07.2011	W. Dubitzky	Initial version of document.
14.08.2012	M. Schiffers	Updates in several areas
29.08.2012	W. Dubitzky	Full revisions
13.9.2012	M. Schiffers	Finalized draft for QAB
21.9.2012	Internal reviewers	Comments and remarks
24.9.2012	M. Schiffers	Final version after integration of QAB comments

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Executive summary

Multiscale modeling and simulation (**MMS**) decomposes a complex problem by calculating properties or system behaviour using information or models from different levels. On each level, a particular approach is employed to describe the phenomenon over a specific scale of length and time. MMS solutions typically involve non-trivial software and hardware requirements. Many state-of-the-art MMS applications address these requirements in an *ad hoc* fashion. The gap between the need to develop, deploy and execute MMS applications on one hand, and the lack of a suitable MMS computing framework is tackled by the MAPPER project. MAPPER's vision is to close this gap through a new computing framework referred to as **distributed multiscale computing** (**DMC**).

The aim of the MAPPER project is to develop concepts, methods and technology facilitating the development, deployment and execution of MMS applications. To ensure that the technology developed by MAPPER will be adopted and further developed after the end of the project, the partners have developed a comprehensive Sustainability Plan.

MAPPER's approach to sustainability is structured into the four layers of tasks and activities. Each layer is associated with various standardization and sustainability tasks, activities and the relevant communities. The four layers are: *science* (science communities employing MMS applications); *ICT R&D* (computer science communities developing new MMS concepts and methods); *software* (ICT developers concerned with high-level and low-level software components, tools and services); and *resources* (e-infrastructures providing the ICT resources for MMS applications).

This document is the final version of the Sustainability Plan. It summarizes the rationale of MAPPER's Sustainability Plan. It (a) provides an update of the dimensions relevant to sustainability, (b) discusses initial outcomes resulting from the implementation of the Plan, and (c) outlines future implementation activities. This document also presents an overview of our thinking in terms of potential future projects in this area. In particular, Appendix B (List of MAPPER Components to be Sustained) has been added.

The previous version of the MAPPER Sustainability Plan and other supporting documents are available through the project website at http://www.mapper-project.eu. Of particular interest in the context of this Final Sustainability Plan are the following documents (short titles):

- D3.1: Report on the Policy Framework for Resource Providers [MAP10a].
- D3.2: Standardization Roadmap and First Sustainability Plan [MAP11d].
- D3.3: MAPPER Profile [MAP11e].
- D3.5: MAPPER Test Suite [MAP12a].
- D3.4: Foresight Study on MMS on e-infrastructures [MAP12b].

1 Introduction

The declared goal of MAPPER is to facilitate a more systematic approach to developing, deploying and executing MMS applications on emerging and future e-infrastructures. One of the key assumptions is that such applications will increasingly be operated in distributed computing environments. Hence, we refer to this type of computing as distributed multiscale computing (DMC).

With this goal in mind and based on a selection of MMS applications from various communities, the MAPPER project is developing novel DMC concepts, methods and tools. Key to the success of the MAPPER project is the sustainability of the developed technology beyond the lifetime of the project. The MAPPER sustainability approach is characterized by the dimensions depicted in Figure 1.

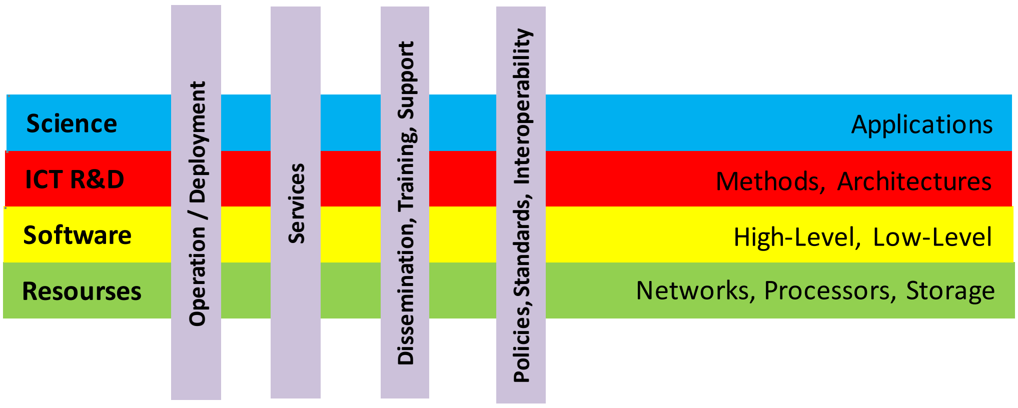

Figure 1: Summary of MAPPER Sustainability Matrix. Rows: Different aspects of science and engineering, both in computer science (*ICT R&D*, *software*, *resources*) and in scientific areas that deploy and use MMS applications (*science*). Columns: Tasks, activities and procedures related to the different aspects of sustainability.

The MAPPER Sustainability Plan has been developed taking into account a Standardization Roadmap (see Appendix A of Deliverable 3.2 [MAP11d]), which describes the standards relevant to on-going MAPPER implementations and mid-term to long-term standardization issues in relation to emerging and future e-infrastructures capable of running DMC applications. The Sustainability Plan and the Standardization Roadmap are complemented by various MAPPER deliverables

- D3.3 (MAPPER Profile [MAP11e]),
- D3.5 (MAPPER Test Suite [MAP12a]),
- D3.1 (Policy Framework for Resource Providers [MAP10a]), and
- D3.4 (Foresight Study on MMS on E-Infrastructures [MAP12b]).

1.1 Deliverable D3.2 Standardization Roadmap

To achieve a lasting impact, MAPPER services and applications need to be deployable on several existing (PRACE, EGI, etc.) and future e-infrastructures. MAPPER, therefore, aims to adhere to and contribute to relevant standards.

The starting point for deriving the MAPPER Standardization Roadmap is represented by the MAPPER software stack as depicted Figure 2.

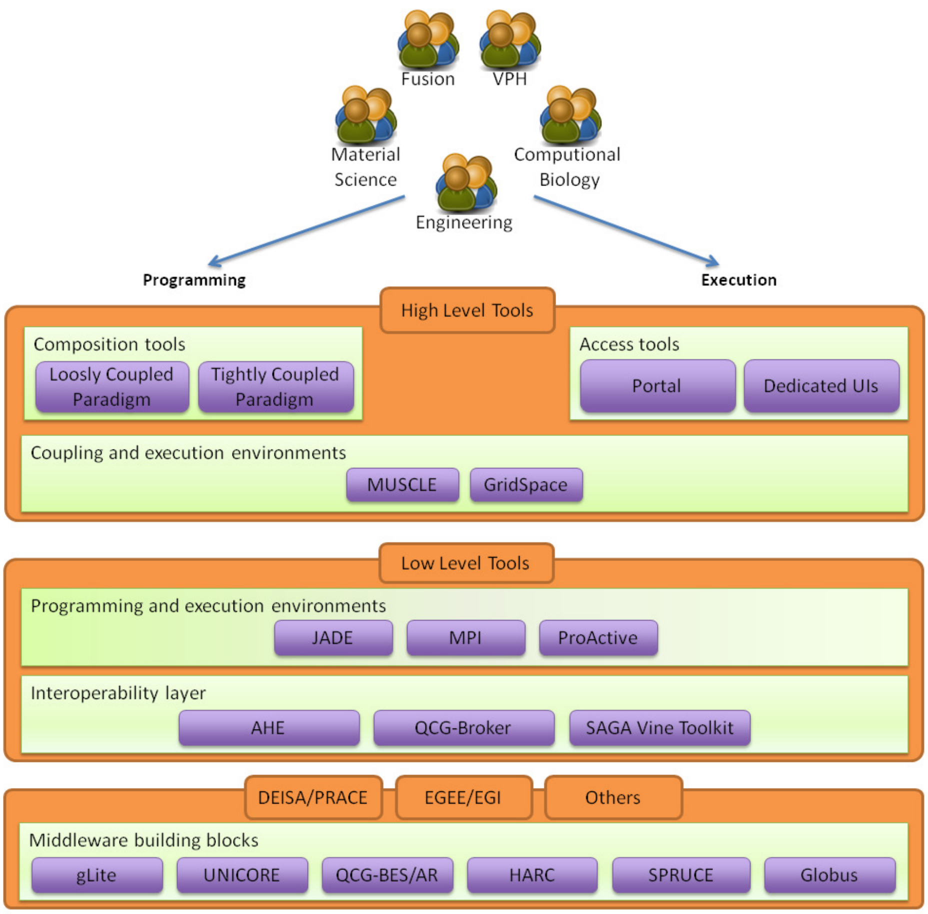

Figure 2: MAPPER software stack (MAP11a).

From a long-term perspective, the MAPPER approach to standardization/sustainability as depicted in Figure 1 represents just a particular instance of the more abstract stack shown in Figure 3. The MAPPER Standardization Roadmap implies an inventory of standards for the interfaces S1, S2 and S3 (Figure 3), serving as starting point for the Roadmap towards a standards-based for DMC. The model is based on a layered approach with (heterogeneous) e-infrastructures on the bottom, a basic programming and interoperability layer in the middle, and a high-level programming and access layer on top. Access to these layers is via the interfaces S1 to S3.

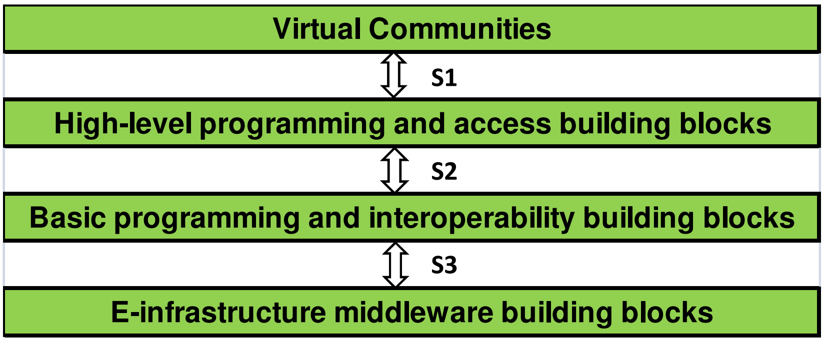

Figure 3: Abstract MAPPER stack.

In order to standardize the interfaces S1, S2 and S3, the following questions need to be addressed.

1. What information schemas are required?
2. What protocol interfaces (messages) are required?
3. What APIs are required?

In order to answer these questions, we require a thorough understanding of current standardization efforts. The MAPPER Standardization Roadmap [MAP11d] and its associated deliverables (MAPPER Profile [MAP11e] and Test Suite [MAP12a]) provide a detailed discussion on the relevant issues which are omitted here for the sake of brevity.

1.2 Deliverable D3.3 (MAPPER Profile)

The basic MAPPER Profile describes a set of requirements, policies and standards that a resource provider needs to implement in order to claim to "MAPPER-compliance". The requirements for the MAPPER Profile are derived from two main MAPPER scenarios; the

1. Loosely coupled application scenario, and
2. Tightly coupled application scenario.

Both scenarios have in common the co-allocation of heterogeneous resources hosted by providers. However, from the point of view of a single service provider, the two scenarios reduce to the simpler MAPPER Multiscale Base Case (MBC). The MAPPER Profile describes the requirements a service provider needs to meet in order to support the MBC directly, and thus, indirectly the loosely coupled and tightly coupled application scenarios. The MAPPER Profile also describes the lower level Open Grid Forum (OGF) standard Distributed Resource Management Application API (DRMAA), how this is exploited by MAPPER components, and how other technology providers could offer services compliant with the MAPPER Profile.

1.3 Deliverable D3.5 (MAPPER Test Suite)

The purpose of the MAPPER test suite is to provide a framework for examining a multiscale system under test (SUT) against a set of objectives (the test cases) expressed by the MAPPER Profile. As MAPPER applications centre on multiscale solutions provided on computational and storage resources, the coordinated and secure access to data and computational resources is an important requirement. Consequently, the MAPPER Test Suite instantiates the generic SUT pattern by providing mechanisms for testing multiscale grid applications against required MAPPER Profile standards.

The MAPPER Test Suite itself is not a specific development. Rather, it is primarily a methodology leveraging the tools provided by the various development environments. It has been setup with the focus on SUTs in three categories: basic test units for evaluating single methods, grid tests for evaluating MAPPER applications within a complete grid setup, and integration tests for examining the interfaces and the interactions among the different subsystems in both the MAPPER deep track and fast track.

1.4 Deliverable D3.4: Future Studies on MMS on e-infrastructures.

As part of the MAPPER project, we are conducting a foresight study on future developments in the field of multiscale modelling and simulation. The aim of this foresight study is to inform policy makers, resource providers, scientists, developers, and other practitioners and stakeholders in this area to enable them to create, anticipate and manage change. In the first stage of this consultation, we ask a selected group of experts to express their views on this subject. Based on the analysis of the feedback, we will generate a more structured and targeted questionnaire which will be used to consult a larger group of stakeholders. Currently, we are still gathering data for the first draft version of D3.4. The final version of D3.4 will be available by the end of the project.

2 Methodology for the Sustainability Plan

The objective of the MAPPER Sustainability Plan is to promote the sustainability of the technology and applications developed during the MAPPER project beyond the end of the project. The sustainability of the MAPPER solutions will depend on the long-term availability of MAPPER components, services and infrastructures; on long-term standards; and on long-term accounting and billing mechanisms. Thus, sustainability is concerned with ensuring (a) that MAPPER solutions will continue to be available and functional, (b) that future developments of MAPPER technology are facilitated and supported, and (c) that training and troubleshooting for MAPPER technologies will be provided.

The methodology for developing the MAPPER Sustainability Plan is based on an analysis of the e-IRG Roadmap 2010 [IRG10] as it describes the fundamental contribution of research e-infrastructures to Europe’s competitiveness. The MAPPER Sustainability Plan follows a similar approach by considering the following layers:

- ***Resources*** (network resources, computational resources, storage resources),
- ***Software*** with low-level basic services and high-level compound services building on the resource layer,
- ***ICT Research and Development*** comprising methods and architectures
- ***Science*** devoted to scientific applications as considered in MAPPER work packages WP7 and WP8.

Figure 1 and Figure 4 depict these layers in relation to training, support, resource and service deployment, tools, and resource and service operation. The full MAPPER Sustainability Matrix is depicted in Figure 4.

		Operation / Deployment		Services	Dissemination, Training, Support			Policies, Standards, Interoperability	
		Independent Provider	VRC Sites	Tools Deployment	Dissemination	Training	Support	Policies	Standards & Interoperability
Science	WP7	VRC-driven, new curricula teaching science from an e-science perspective. MAPPER supports depoyment/usage of multi-scale modeling and simulation tools, including MAPPER tools.							
		WP6	WP4 / WP5	WP4 / WP5 /WP7	WP2	WP2	WP4 / WP8	WP3	WP3
ICT R&D	Methods Architecture	-	-	-	- School - Publications - WS/conference	School	-	-	-
Software	High level	-	Target specific VRCs: fusion, VPH, biology	- Sustained by partners	- D2.2: Diss. Plan - D2.4n: Diss. Reps - Publications - WS/conference	- School - Training materials	Yes* (at a cost)	- Support MAPPER-related VRC - EGI vs PRACE	- Use standards - Use OS-agnostic software
	Low level	EGI/PRACE Evaluation proc.	See deliverable D5.1	-			Yes**.	- D3.1 → e-IRG, Comm. agencies	- Yes, support using OGF - Active involvement OASIS
Resources	Network Cycles Storage	-	Statistics in D6.2.	-	Statistics published in papers, on website, etc.	-	-	- Allocation policies - Resource allocation procedures - SLA, gSLA - e-IRG, GÉANT - Policy management board - EGI vs. PRACE policy	- Allocation standards - Collect best practises - D6.1 w.r.t. security

* Support is provided by MAPPER partners through e.g. project membership. We are considering to adopt other business models. In addition, we plan to integrate our software with the EGI infrastructure.

** Most of the software will be available as open source. The QCG layer is supported by PSNC. See Appendix B.

Key:

WP1	**Management**	VRC:	Virtual research communities	e-IRG:	e-Infrastructure Reflection Group
WP2	**Dissemination and Outreach**	Dess. Plan:	Dissemination Plan	OGF:	Open Grid Forum
WP3	**Policy Support and Sustainabil**	School:	MAPPER School	OASIS:	Open Standards for Info Society
WP4	**Adaptation of Existing Services**	WS:	Workshop	EGI:	Open Grid Infrastructure
WP5	**Vertical Integration**	OS:	Operating system	PRACE:	Advanced Computing in Europe
WP6	**E-Infrastructure Operations**				
WP7	**Application Enabling and Evaluation**				
WP8	**Programming and Execution Tools**				

Figure 4: Diagrammatic summary of the MAPPER Sustainability Plan.

3 Science

Currently, the main type of applications that use MAPPER technology comes from science (and engineering) domains. Hence, a key part of MAPPER's approach to sustainability is to ensure that MAPPER concepts and solutions are used by a wide range of science and engineering users that require DMC capabilities. Initially, MAPPER is focusing on a small set of application areas including biomedical and nano-material science, hydrology and flow control, and nuclear fusion. The principal approach consists of facilitating support and training for deploying and using DMC tools, including MAPPER tools, for scientific investigations in these areas. Among other things, this may be achieved by attracting funding for new research projects in these areas. An additional element of sustainability is concerned with the development of new curricula teaching e-science, in particular with emphasis on MMS and DMC.

Below we briefly outline aspects of sustainability in relation to the main scientific areas covered by the project. For more details we refer to [MAP11d].

3.1 HemeLB (VPH Community)

The work performed on HemeLB [Maz08] within MAPPER is a starting point for a sustained evolution of HemeLB into a large multiscale method to resolve blood flow in human arteries. The emphasis on the code is to provide medically valid results, practical clinical assistance, and real-time visualisation. HemeLB has already been in use at UCL since 2007 and currently being enhanced.

MAPPER introduces a hierarchical coupling of scales to efficiently model blood flow from the main artery scale all the way down to the capillaries. To ensure the sustainability of this tool, MAPPER is pursuing several directions. First of all, MAPPER aims at validating HemeLB for use in clinical environments through extensive testing in collaboration with the UCL Hospital (UCLH). Second, MAPPER will incorporate the resulting improvements directly into the HemeLB codebase. As a result, MAPPER will be able to reuse the results in several large projects which feature HemeLB as an example application. First, HemeLB will be part of the 3-year FP7 project CRESTA (Collaborative Research into Exascale Systemware, Tools and Applications), which started in October 2011. Second, the code will be central in the EPSRC-funded project LSLBBC (Large Scale Lattice-Boltzmann simulation of BioColloids), which began in January 2012. This particular project will fund a research at UCL and will involve enhancement of HemeLB and subsequent use. And third, the code will be adopted within a 3-year 2020 science project which started in July 2011.

3.2 Material Science

The rationale for sustaining multiscale approaches in MAPPER's nano-material applications focuses on adopting multiscale approaches to enable the modelling and simulation of more complex systems and to obtain a more robust statistical sampling of the properties of clay-polymer interactions.

The main sustainability driver in the nano-material domain is the adoption of widely used and open-source tools. All simulation codes used within the nanomaterial application (LAMMPS[1], CPMD[2]) are publicly available and are coupled using core MAPPER components such as GridSpace, AHE and QCG-Broker. In addition to ensuring that all components are open-source, MAPPER aims at including user manuals for deploying and operating multiscale simulations using these codes. Although MAPPER will provide manuals and publish its methods, MAPPER will not be providing support on an individual user basis to communities.

Additionally, MAPPER intends to apply its multiscale solutions in a project from the Qatar National Research Fund, named "From Fundamental Understanding to Predictive Design of Layered Nanomaterials", which began in December 2010 and which will run for 3 years.

3.3 In-stent Restenosis and Cranial Aneurysms

The multiscale nature of biomedical systems is very well recognized in the physiology communities, as well as the need for consistent approaches to the coupling of sub-models where many of them need High Performance Computing (HPC) facilities to be executed (see VPH framework). Within the EU-funded projects MeDDiCA[3] and THROMBUS[4], the UvA is concerned with the modelling of In-stent restenosis and stent-induced thrombosis in cranial aneurysms.

---

[1] http://lammps.sandia.gov/

[2] http://www.cpmd.org/

[3] http://www.meddica.eu/

[4] http://www.thrombus-vph.eu/

THROMBUS is an EU-funded FP7 collaboration of biomedical and computational science research institutions, hospitals, and a stent company that started in February 2011 and will run for 3 years. THROMBUS is concerned with the development and validation of a multiscale model for the stent-induced thrombosis in cranial aneurysms. Besides a better understanding of the processes involved in thrombosis, the ultimate goal is to develop a multiscale computational model and simulation framework that can be used in clinical environments to support clinicians in choosing the right stent and the most effective deployment strategy. Large-scale distributed multiscale computation facilities enable both, the thorough validation of a simplified multiscale model that can run locally in clinical environments, and the execution of the large-scale multiscale model for patient-specific data in an urgent computing scheme. The methodologies developed within MAPPER will be applied. MAPPER activities regarding policies and computing services are important steps to enable urgent biomedical multiscale computing in the future. The THROMBUS project will have direct impact on more predictive, effective and safer healthcare.

MeDDiCA is a Marie Curie Initial Training Network (MC ITN) focused on Cardiovascular Engineering and Medical Devices that started in September 2009 and will run for 4 years. Within MeDDiCA at the UvA we will further develop a computational multiscale model for in-stent restenosis (ISR). The ISR model has been developed during the FP6 COAST project that lay the foundation for many aspects of multiscale modelling and execution now brought to a next stage within MAPPER. The ISR model is one of the most advanced use cases in MAPPER and MAPPER tools will be directly implemented and validated using this model. Main objective within MeDiCCA is to improve the existing ISR model and enrich its predictive capabilities regarding the abnormal growth of smooth muscle cell tissue after stenting arteriosclerotic arteries. A number of biomedical, clinical, and computational science research institutions are involved in this project, as well as a simulation software company, all working on the modelling and understanding of cardiovascular diseases and the design of cardiovascular devices. Within this framework multiscale methodologies and simulation services developed within MAPPER are actively promoted and shared with the other partners.

The methodologies, software frameworks, and services developed within MAPPER facilitate the development and execution of these models and will be further promoted through the projects' dissemination activities and subsequent research collaborations.

Through our involvement in the VPH community MAPPER methodologies and tools will eventually also find its way into the VPH Toolkit, an intended extensive collection of biological sub-models and software frameworks to couple them.

More in general, the European wide VPH community has organized themselves into the VPH Institute. UvA and UCL are members of the institute and through the institute dissemination of MAPPER tools and services will be undertaken. Moreover, MAPPER will establish and foster contacts between the VPH institute and other e-infrastructure projects such as PRACE[5] and EGI-InSPIRE[6], thus fostering further dissemination of MAPPER results.

3.4 Flow Control

Currently, the MAPPER project offers a promising framework to develop large multiscale applications. UNIGE is involved in the simulation of irrigation canals and other water courses. Flow control is an important challenge for computational science. Many practical problems require us to act on some part of a fluid flow. This is the case of the management of irrigation canal where gates should be adjusted dynamically to fulfil the water demand, as a function of the meteorological situation or the need of the users. Obvious constraints are to avoid the overflow or draining of the canal, or a waste of water. Other flow control problems are related to the accumulation of sediments in lakes, near canal gates, etc. By adjusting flow parameters, the deposit can be brought to adequate places.

In a completely different context, blood flow control is becoming of central importance in many biomedical application (see also in-stent restenosis above), and in particular in cerebral aneurysms. Devices, called flow diverters (or stents) are inserted in the parent artery to control the flow in the aneurysm cavity and produce its occlusion.

3.5 Computational Biology

The University of Ulster focuses on MMS applications in the fast-emerging interdisciplinary area of systems biology. While many approaches to systems biology involve single-scale models, a growing body in recent research aims at modelling life

[5] www.prace-project.eu

[6] http://www.egi.eu/about/egi-inspire/

phenomena across several scales [Men11] involving multi-scale/multi-model approaches. Multiscale systems biology is concerned with experiments and hypotheses that involve different scales of biological organization from intracellular molecular interactions to cellular behaviour and the behaviour of cell populations, to an organ or even a whole organism. While many models in systems biology require solutions that deal biological processes that at multiple space and/or time scales or different levels of biological organization, there is also a need to quantitatively explore biological systems by means of simulations that incorporate heterogeneous DMC techniques, and to scale the computation of simulations. To facilitate a long-term sustainability of MAPPER technology, UU focuses on standard, widely used computational biology tools and technology, including the Systems Biology Markup Language (SBML), Systems Biology Results Markup Language (SBRML) and COPASI [Hoo06].

Systems biology is an extremely broad field involving a large number of diverse biological phenomena. The focus of the Computational Biology application is on modelling and simulation of gene-regulatory networks (GRN). The process of gene regulation is essential for life as it increases the versatility and adaptability of an organism by allowing it to express protein when needed. To date, the majority of gene regulation models address GRN systems that contain less than 20 genes. The reason for this is that the computing requirements for reverse-engineering such models from data grows exponentially with the number of genes. The computing requirements are also dependent on the connectivity topology of the underlying GRN system, the number of parameters in the chosen GRN rate law, the given time resolution, the numerical integration method chosen, and other factors. Currently, there is a trend in biology to investigate increasingly larger GRN systems; ultimately many important life processes are regulated by more than 20 genes (see, for instance, the study by Davidson & Levin [Dav05]). Such large GRN systems are characterized by an inherent biological organization composed of sub-networks, each regulating a separate biological function or process while still interacting with each other. The sub-networks in such a multi-network structure may operate on distinct temporal scales. Applying current GRN modelling and simulation approaches to such large, multi-network GRN systems is problematic, both conceptually and computationally. A promising approach to address these challenges is MMS. In the MAPPER

Computational Biology application, our aim is to explore this area further and provide tools facilitating research in this area.

4 ICT R&D

Sustainability of ICT R&D within MAPPER concerns the scientific methods and architectures, which are developed within the project. The outputs of these activities will be sustained by disseminating methods and architectures as applied in MAPPER to a wider community. This will be accomplished by training activities, technical documentation, and high-quality publications for peer-reviewed scientific journals and conferences.

Training

Within the lifetime of the project, MAPPER organizes at least two MAPPER Seasonal Schools. Their aim is to provide training on MAPPER concepts and tools. The First Seasonal School took place at UCL in February 2012. It attracted 25 participants with 36% from outside the project. Several external speakers[7] highlighted the importance of MMS. The second MAPPER Seasonal School, scheduled for year three, will be fully open to researchers, developers, scientists and other interested parties. Once the MAPPER tools are deployed on the EGI service stack (and later on the PRACE stack), the School will also be advertised via EGI and PRACE promotional channels.

Publications and presentations

MAPPER members aim to regularly attend a wide range of scientific conferences and to publish scientific methods and architectures in international journals and conference/workshop proceedings. The MAPPER Dissemination Plan [MAP10] provides a detailed description of the publication approach.

Community building

Because DMC is an emerging field, MAPPER sustainability is not only focused on effective dissemination and training, but also on building a sustainable community for

[7] **Daniel S. Katz**, Senior Fellow in the Computation Institute at the University of Chicago and Argonne National Laboratory, USA; **Gabor Z. Terstyanszky**, Reader in Computer Science and Software Engineering at the University of Westminster, UK; **Ivan Kondov**, Teamleader SimLab NanoMikro in Performance assessment of different constraining potentials in computational multiscale materials and biomolecular simulations at Simulation Lab NanoMikro, Karlsruhe Institute of Technology, Germany; and **Mihai Duta**, Scientific Software Advisor for the Oxford Supercomputing Centre, UK.

DMC. By contributing to an MMS R&D community, MAPPER aims at increasing the public awareness of DMC.

MAPPER's initial effort into bolstering an interdisciplinary multiscale community is based on organising international workshops for Distributed Multiscale Computing (like the one at the IEEE e-Science conference[8] which took place in December 2011), both during and beyond the project's lifetime.

Because community building also relates to co-operation with similar projects, MAPPER supports related EU-funded projects like DRIHM[9] and upcoming the lifetimes of which will exceed that of MAPPER.

5 Software

The process of creating MAPPER software and integrating it with existing e-infrastructures is based on predefined reusable software components (Figure 2). The MAPPER consortium has contributed a number of legacy software components that need to be integrated and maintained within the e-infrastructures to support new or improved capabilities offered to users and their DMC applications. From the sustainability perspective, it is important to understand that MAPPER distinguishes two software development phases, *fast-track* and *deep-track*, respectively. In order to ensure early adaptation, integration and deployment of a minimal set of infrastructure components to facilitate the coupling of multiscale applications, the fast track approach was implemented. The deep track, on the other hand, will be responsible for a more formal adaption, integration and deployment of higher level services required to realise the fully integrated MAPPER infrastructure which will facilitate the automatic coupling and launching of multiscale components. While MAPPER has already reused available software components and integrated them into a consistent middleware (based for example on QosCosGrid v2.0 (Figure 6)) during the fast-track phase, the deep track approach is more relevant to sustainability as it goes a step further and provides automation mechanisms based on reusability and standards.

By *reusable component* we do not only refer to code, but also MAPPER software specifications, reference implementations of standards, application use cases or

---

[8] http://www.computationalscience.nl/dmc2011/

[9] http://www.drihm.eu/

templates. The fast-track components deliver basic functionality needed, in particular, for tightly coupled MMS applications that need to be executed and managed on co-allocated computational resources across many sites in the e-infrastructure and their (quasi-)standardized software stacks.

From a sustainability and maintenance perspective, MAPPER components fall into four groups:

1. Local domain (low level) MAPPER services like QCG-Computing (an advance reservation and execution service) or QCG-Notification (a notification service).
2. Grid domain (high level) MAPPER services and user interfaces like QCG-Broker (a load balancing and co-allocation service) or Application Hosting Environment (AHE) which facilitates application specific services to utilize Grid resources.
3. Multiscale coupling libraries like MUSCLE (a tool for tightly coupled MMS applications) or GridSpace for facilitating loosely coupled MMS applications.
4. Third party e-infrastructure services maintained and supported by external communities (e.g., UNICORE, gLite or Globus Toolkit)

The first group of software components belongs to the fast-track phase, whereas the other two groups belong to the deep-track phase. Third party e-infrastructure services are out of scope of the MAPPER Sustainability Plan as external communities support them directly. MAPPER has delivered their software (both fast track and deep track) to EGI and PRACE for deployment evaluation – a significant step forward to sustainability.

The deep-track software creation and integration phase has stared in the middle of the project. It should be noted that high-level services and application tools of the deep-track phase are not directly integrated with e-infrastructure services and less efforts is required to support the software as it will be deployed in the user space. Once the access to new capabilities in MAPPER is provided to end users, they can modify and maintain their own software depending on their requirements.

Similarly to resources, sustainability of software also needs to be considered from different angles: from an operational perspective, a service perspective, an educational perspective, and an interoperability perspective according to Figure 4. In the next subsections we will address these aspects of sustainability by describing various strategies reflecting different classes of software components.

5.1 Operation and Deployment View

This section briefly describes the MAPPER software from both operation and deployment perspectives. The common feature of middleware services is that they must be deployed and then supported locally by a system administrator of a certain e-infrastructure site; whereas application tools can be deployed by users that have authorized access to the site (see Table 1).

Table 1: Software capabilities.

Software	Capabilities						
Middleware Services	**Basic Jobs**	**MPI Jobs**	**Loosely Coupled Jobs**	**Tightly Coupled Jobs**	**Advance Reservation Support**	**Automated Notifications About Jobs**	**Automated Co-allocation Support**
QCG- Computing	x	x	x	x	x		
QCG-Notification						x	
QCG-Broker	x	x	x	x	x	x	x
AHE	x	x	x		x		
UNICORE	x	x	x				
gLite	x		x				
Application Tools							
GridSpace	x	x	x		x	x	x
MUSCLE		x		x	x		x
SAGA / Vine Toolkit	x	x	x		x		

Some MAPPER software components may require additional privileges (e.g., the job submission and advance reservation service). Therefore, further deployment and administrative efforts may be required at various sites. Detailed configuration descriptions, templates and examples are therefore provided after a successful deployment evaluation based on the objectives expressed in close cooperation with EGI and PRACE[10].

An example is the security audit because most e-infrastructures shape their own policies and procedures for software installation and security policies and most

[10] The activities are coordinated by the MAPPER-PRACE-EGI Task Force (https://wiki.egi.eu/wiki/MAPPER-PRACE-EGI_Task_Force_%28MTF%29)

MAPPER components already passed such procedures in National Grid Initiatives (NGI).

5.2 Services view

Within the Service Activity work packages, MAPPER has published Deliverable D6.1: Report on the Assessment of Operational Procedures and Definition of the MAPPER Operational Model [MAP11b] which is concerned with the services view on MAPPER software. The report discusses the assessment of operational procedures of the European e-infrastructures targeted by MAPPER, including DEISA, EGI and PRACE, and defining the operational model for MAPPER software.

5.3 Dissemination, training and support view

The software employed in the MAPPER project benefits from various software components provided by the consortium which have been developed and supported over the last few years (Figure 6). As a result, the project partners were able to setup core middleware services, including QCG-Computing, QCG-Notification and QCG-Broker, on selected production EGI sites (e.g., the Polish PL-Grid and the German NGI) early in the project. This effort has a direct impact on other NGIs in EGI and resulted in a task force within EGI to support MAPPER. Thus, some of dissemination, training and support procedures will be adopted by MAPPER in the near future.

Additionally, many training and dissemination activities for different software components have been already performed under the EU funded projects:

- **QosCosGrid**: developed since 2005, supported in QosCosGrid, BREIN and PL-grid projects, disseminated during various conferences talks, interoperability events and publications. Currently the QosCosGrid stack is deployed on all PL-Grid production sites.
- **GridSpace**: developed since 2006, supported in Virolab, Gredia, PL-Grid projects, disseminated in tutorials held at the conferences, several conference and journal publications, with ongoing training activities addressed to Polish scientific communities.
- **AHE**: developed since 2005, supported in Virolab, OMII-Europe and RealityGrid projects.

The contracted support within the PL-Grid project covers a 5-year period ending in 2017. Until then, both GridSpace and QosCosGrid will be available for end users (including support). As indicated, the original components of the MAPPER project have been and will be disseminated and supported (including user support). On the other hand, the compilation of already existing original elements enhanced with specific MMS applications' demands need to be addressed as a whole and presented to MMS applications' communities as a comprehensive environment dedicated to them. In any case is the sound and solid foundation of the consortium partner institutions and their obligations to support the software and the users in the next several years as well as the already performed sustainability and dissemination efforts related to the MAPPER software a good path towards sustainability.

5.4 Policies view

From the MAPPER software perspective, especially for many-cluster multiscale model executions, the most crucial policies are site networking policies that concern network firewall configuration. Those network policies usually differ between sites. However, a minimal set of requirements is the following:

- All traffic is allowed between infrastructure nodes.
- Interactive nodes allow outgoing connections to external sites.
- Subject to special agreements some inbound ports can be opened on interactive nodes.

For this reason, all communication and networking channels established between single-scale modules located on different e-infrastructure Grid/HPC sites need to be routed via transport overlay deployed on interactive nodes. The detailed description of networking policies is reported in the D5.2 MAPPER Vertical Integration Plan [MAP11f].

5.5 Standards and interoperability view

Many Grid/HPC middleware technologies have been augmented with implementations of proposed standards from the Open Grid Forum (Figure 6). From the middleware interoperability perspective, the key specification is the OGSA Basic Execution Service (BES) for job management and submission. By supporting the BES, it is possible to improve the interoperability between different Grids or HPC environments as well as to allow upper-level services and application tools, e.g. SAGA, Vine Toolkit or GridSpace, to provide efficient access to remote computational resources in a uniform manner. BES provides easy, intuitive and standardised access to computational resources. The BES specification defines two mandatory (BESFactory and BESManagement) and one optional (BESActivity) interface. The management interface allows controlling the service itself. The factory interface provides job submission and bulk monitoring capabilities while the activity interface allows monitoring of a single job. All jobs submitted to the BES interface have to be described in the Job Submission and Description Language (JSDL) (an OGF standard relevant to MAPPER from interoperability perspective[11]).

Efforts are required at the high-level application programming layer to bridge the gap between existing Grid middleware and application-level needs. The Simple API for Grid Applications (SAGA) is an OGF standardization effort that addresses this particular gap by providing a simple, stable and uniform programming interface that integrates the most common grid programming abstractions[12]. These abstractions have been identified through the analysis of several existing and emerging applications, including MMS application scenarios. As presented in Figure 6, all relevant Grid middleware services commonly used in the e-infrastructure, including gLite, UNICORE, Globus Toolkit and more importantly QosCosGrid, support the aforementioned standards for interoperability.

As a proof of this sustainability concept, MAPPER has already performed interoperability and performance tests to demonstrate remote job submission and monitoring provided by the most recent stable and recommended Grid middleware services for e-infrastructures, namely gLite 3.2, UNICORE 6.4, Globus Toolkit 5.0.3 and QosCosGrid 2.0. During these test it was shown that JSDL defined jobs could be submitted via high-level SAGA libraries to underlying Grid middleware services via BES interfaces. As a result (see Figure 5), the QosCosGrid middleware outperformed other available Grid middleware services by almost an order of magnitude. Similar throughput benchmarks confirmed that the QosCosGrid middleware adopted in MAPPER is the most efficient implementation of interoperability standards. All

---

[11] http://www.ogf.org/gf/group_info/view.php?group=jsdl-wg

[12] http://www.ogf.org/gf/group_info/view.php?group=saga-wg

benchmarks have been performed on selected EGI production sites available for MAPPER users. Detailed descriptions will be included in appropriate Service Activity deliverables and in technical reports submitted to OGF.

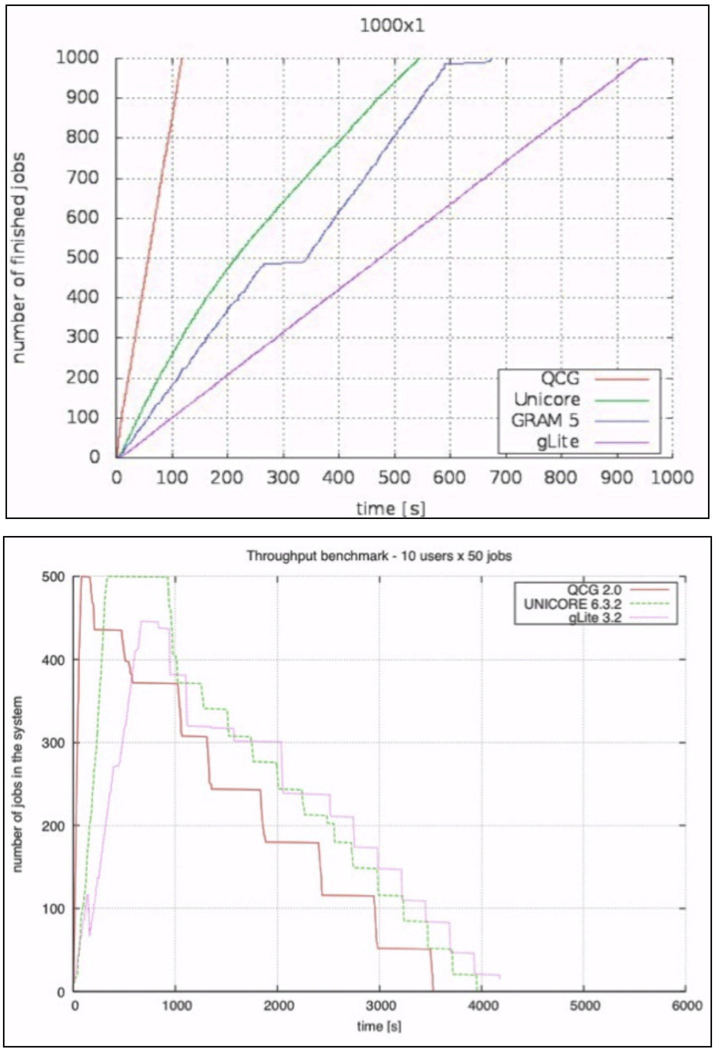

Figure 5: The overall performance of the QosCosGrid BES implementation compared to UNICORE, gLite, and Globus.

We envision that further interoperability improvements will include tests of new capabilities of the e-infrastructure, e.g., advance reservation defined by new versions of the OGF DRMAA standard[13]. Currently, however, is QosCosGrid the only Grid middleware supporting advance reservation and co-allocation.

[13] http://www.ogf.org/gf/group_info/view.php?group=drmaa-wg

A concluding result as far as sustainability planning is concerned, is that there are still many efforts needed to improve the interoperability between components and their heterogeneous policies (e.g., authentication and authorization). A long-term solution is beyond the scope of this report.

Supporting long-term solutions, however, MAPPER (by its WP3 activities) will use and further promote standards on many different layers. The planned actions will focus on:

- active participation in standardization bodies, in particular OGF and OASIS,
- adopting new standards in MAPPER components, such as OGF DRMAA, OASIS SAML2.0, etc.,
- providing feedback on evolving standards,
- lobbying on deploying software components that adhere to open standards,
- taking part in interoperability events, like the OGF BES interoperability demo[14].

Finally, as the general rule, all software components delivered by MAPPER will be OS-agnostic software and will follow the commonly accepted software packaging guidance[15].

6 Resources

Following Foster et al. [Fos01], (computing) *resources* in the broader sense of this deliverable are physical resources like network elements, computing elements and storage elements. Such computing resources are “owned” by *resource providers* (i.e., organisations which are legal entities) and provided to Virtual Organisation (VO) for temporary use, subject to policies and constraints. Consequently, the perceived (virtual) life cycle of resources is coupled to the life cycle of VOs and VO membership of resource users. After a VO is decommissioned, the resources are no longer available to the members, unless they join another VO which grants new access rights. VOs represent structures implemented in the e-infrastructure that facilitate access to resources that span multiple administrative and geographical locations. Virtual Research Communities (VRC) are collaborations of “like-minded” individuals

[14] Interoperability Event at OGF30: https://forge.ogf.org/sf/go/wiki2458

[15] Fedora Packaging:Guidelines: http://fedoraproject.org/wiki/Packaging:Guidelines

that work in the same discipline, share a common interest or computational model which may span multiple VOs [IRG10]. VOs are commonly created and managed within VRCs.

MAPPER MMS applications inherently rely on variety of resources of from different providers for various VOs in various VRCs [MAP11a, IRG10]. Examples include HPC systems provided through the Partnership for Advanced Computing in Europe (PRACE), the data communications infrastructure provided by DANTE[16], and the storage capabilities available through EGI and the NGIs.

Because MAPPER – as a project – is not designed to provide any resources to VOs, MAPPER is completely decoupled from resource provisioning by respective resource providers. On the other hand, the intention of MAPPER is to sustainably provide multiscale services. Consequently, arrangements have to be made on the resource level for enabling long-term access to required resources, both for members of VOs (whose life span outlives the MAPPER project) and for members of VOs whose establishment is scheduled after the project's end.

Resource sustainability needs to be considered from different angles: from an operational perspective, a service perspective, an educational perspective, and an interoperability perspective according to Figure 4.

6.1 Operational/deployment view

Assessing the sustainability of resources in the operational/deployment view requires differentiating between resources provided by "independent" resource providers (i.e., those who neither contribute to the MAPPER project directly nor indirectly) and resources provided by VRCs associated with MAPPER.

MAPPER will neither provide any resources to VOs nor will MAPPER be allowed to operationally manage resources outside the ownership of MAPPER partners (unless explicitly allowed by the resource owners). Thus, sustaining any resources (of independent resource providers and VRCs) is beyond MAPPER's scope. Despite this general restriction, however, there is a difference between independent resources and VRC resources. While the life cycles of the former are completely independent

[16] http://www.dante.net/

from MAPPER activities, the life cycles of the latter may have a relationship with MAPPER, which is typically externalized as

1. A direct project contribution (e.g., parts of the cluster at the Leibniz Supercomputing Centre);
2. A service-level agreement (SLA) formally specifying the terms and conditions of guaranteed resource provisioning; or
3. A memorandum of understanding (MoU) describing an agreement with the VRC sites (acting as resource providers) for a convergence of will with an intended common line of action related to resource provisioning.

Note that these externalizations are time-dependent: While a direct contribution (bullet 1) is only possible at the lifetime of the project, all other externalizations (bullets 2 and 3) may also be relevant after the project's end.

From a sustainability (of resources) point of view, MAPPER will therefore instigate a series of MoUs and SLAs with resource and infrastructure providers for long-term and dependable resource provisioning. A first MoU has been signed with EGI[17]; MAPPER has initiated also an evaluation process with PRACE. After the first evaluation step MAPPER middleware components have been accepted for testing procedures on selected PRACE sites, in particular SARA in Amsterdam, Netherlands. Long-term cooperation agreements with American initiatives like XSEDE[18] and ExTENCI[19] are just as desirable as with Asian and Australian ones. Depending on the project results, MAPPER may provide as a separate deliverable a sample SLA to be used for after-project resource provisioning. Note, however, that such SLAs need to be associated with cost models, accounting and billing.

In order to simplify the agreements of SLAs and MoUs (within the lifetime of the project and beyond), MAPPER will keep an inventory of all deployed services in the public Deliverable D5.1 [MAP11g]. This inventory may also serve as a checklist for arranging multi-SLAs with different providers. Furthermore, the operational statistics

[17] http://www.mapper-project.eu/documents/10155/25657/EGI-InSPIRE-MOU-MAPPER-FINAL.pdf

[18] https://www.xsede.org/

[19] https://sites.google.com/site/extenci/

collected in Deliverable D6.2 [MAP11c] provide a good basis for structures SLAs and MoUs will typically rely on.

6.2 Service view

Because this section discusses sustainability of (physical) resources, a separate service view is not necessary. Instead, we refer to Section 5 for a more detailed analysis.

6.3 Educational view

The educational view refers to three separate aspects:

- Dissemination and exploitation of results
- Training in using results
- Support while using results

6.3.1 Dissemination and exploitation of results

A sustainable provisioning of resources for MAPPER applications not only requires the resources themselves (see Section 6.1), but the *right* resources. In order to convince resource providers to contribute their resources, they need to know the basic resource requirements. The MAPPER requirements have been reported Deliverable D6.2 [MAP11c].

6.3.2 Training in using results

MAPPER will not provide trainings with a focus on using physical resources. This section is thus not applicable.

6.3.3 Support while using results

MAPPER will not support physical resources. Thus, this section is not applicable.

6.4 Interoperability view

The interoperability view not only covers standards and interoperability issues but also policies. Achieving interoperability is necessary due to the heterogeneity not only of resources, but also of organizational structure and policies. From a sustainability perspective, solving interoperability issues is mandatory when trying to add new resources.

6.4.1 Policies

Regarding the sustainability of resources, the policy aspect of the interoperability view needs to address two questions:

1. How can resources be allocated?
2. Which kinds of SLAs are required?

6.4.1.1 Resource Allocation

Resource allocation policies differ from provider to provider. A generic allocation description is therefore impossible. Hence, we refer to the specific allocation procedures as relevant to the individual providers. In the context of MAPPER, the following procedures are relevant to provide a sustainable allocation of resources. Note, however, that most of these procedures are subject to changes without further notice to MAPPER. Thus, applicants are strongly advised to check the respective latest version.

Resource allocation in GÉANT

GÉANT[20] is the multi-gigabit pan-European research network providing connections between the networks of its European Research and Education Networks (NREN) partners. However, GÉANT connects individual researchers indirectly through a "network of networks". Data sent by one end facility to another one is typically be transferred across several local and regional networks before it reaches GÉANT. GÉANT provides the international connection between the sending end facility and the receiving one. At the receiving end, the data flow through several local and regional networks in the reverse order to that described before.

NREN networks are connected to GÉANT via access links the speed of which depends on the respective NREN subscription, the capacity requirements of the NREN, and the widely differing availability and prices of connectivity across Europe.

DANTE, the coordinator of the GÉANT project (not the network), also offers a so-called DANTE World Service (DWS) for connectivity to the wider Internet.

GÉANT provides inter-country connectivity for Europe's NRENs only, not for end users. From a resource sustainability point of view, allocating a network resource

---

[20] http://archive.geant.net/server/show/nav.159

requires to adhere to the national NREN procedures, as for instance exemplified in the procedures of the German DFN-Verein[21].

Resource allocation in PRACE

In order to assist parties interested in using PRACE[22] resources, the PRACE consortium has prepared a comprehensive Application Guide. It describes in detail the various aspects of preparing a proposal requesting PRACE resources in response to calls for proposals, how the proposal is handled after submission, including peer reviews and resource allocation. The Application Guide is available from the PRACE home page.

Applicants should also be aware that the efficient use of PRACE systems requires detailed knowledge of architecture-specific factors influencing the performance (e.g., compilers, tools, libraries). PRACE, therefore, offers a set of best practice booklets[23] on how to achieve good performance on these systems.

Resource allocation in EGI

EGI[24] is a pan-European Grid infrastructure which was established in close collaboration with the various European national Grid initiatives and the European International Research Organisations (EIRO). This pan-European involvement ensures long-term availability of a generic e-infrastructure for all European research communities and their international collaborators. EGI (or more precisely EGI.eu as EGI's maintainer) does not own any resources; they still belong to and will be managed by the relevant national Grid initiatives. EGI resource allocation is therefore formally externalized by a contract with and adhering to policies of the respective national Grid initiatives.

Since scientific research is expanding worldwide (the e-science paradigm), a longer lasting and easier operational model is needed – both for coordinating the e-infrastructure itself and for delivering integrated services that cross national borders. This is the motivation behind the EGI-InSPIRE project[25] which intends to support the

---

[21] http://www.dfn.de/en/services/dfninternet/

[22] http://www.prace-ri.eu/Application-Guide-and-PRACE-Peer?lang=en

[23] http://www.prace-ri.eu/Best-Practice-Guides?lang=en

[24] http://www.egi.eu/about/

[25] http://www.egi.eu/about/egi-inspire//

transition from project-based systems to a sustainable pan-European e-infrastructure. EGI-InSPIRE will collect user requirements, provide coordination and support for current and emerging user communities, for example, the ESFRI projects[26] (especially those with considerable demands such as high energy physics, computational chemistry, or materials science – application areas fundamental to MAPPER), and define policies and procedures for pan-European resource allocation.

From MAPPER's resource sustainability point of view, these procedures need to be adopted and promoted as soon as they will be available.

6.4.1.2 Service-Level Agreements

Once resource allocation procedures are finalized, MAPPER will provide sample versions of dedicated services-level agreements (SLAs) for various constellations (see also Section 6.1) taking into account similar efforts in EGI.eu, PRACE, DANTE and national bodies (NREN, NGI). The MAPPER SLA samples will also leverage the work performed in the European gSLM project (Service Delivery and Service Level Management in grid Infrastructures)[27], once the project is finished.

6.4.2 Standards and Interoperability

Sustainably allocating, using and managing resources requires dedicated standards and demonstrated compliance with these standards. While most of these standards (or de facto standards) will be defined by the parties mentioned in Section 6.4.1, others are defined by more generic standardization bodies like OGF and OASIS.

[26] http://ec.europa.eu/research/infrastructures/index_en.cfm?pg=esfri

[27] http://gslm.eu/

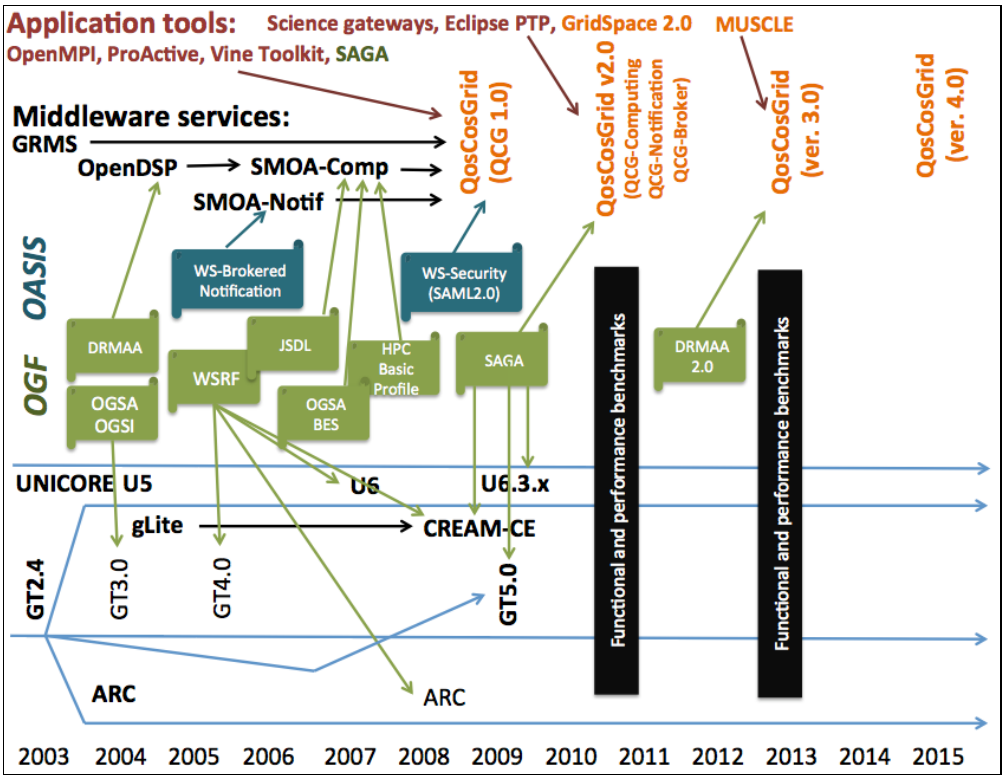

Figure 6: Middleware software evolution (QosCosGrid, UNICORE, gLite, ARC) and relevant OGF and OASIS standards

The most important OGF and OASIS interoperability standards that are supported in MAPPER together with a roadmap are presented in Figure 6. In order to support a sustainable access to resources, MAPPER contributes to several standardization bodies [MAP11d]. MAPPER goes further by defining a specific MAPPER profile [MAP11e] with a corresponding test suite [MAP12a] to sustainably determine the compliance of future applications to the profile.

7 Future projects

Sustainability is closely related to the ability of securing funding for new projects. From MAPPER's perspective, the sustainability drivers in this context are the e-infrastructure Reflection Group (e-IRG[28]) and the European Commission's Horizon 2020[29] framework.

[28] http://www.e-irg.eu/

[29] http://ec.europa.eu/research/horizon2020/index_en.cfm

The e-infrastructure Reflection Group (e-IRG) has provided strategic advice and guidance on the development of European e-infrastructures for science and research for almost a decade. E-IRG's mission has originally been the creation of a policy and administrative framework for the easy and cost-effective shared use of electronic resources across Europe. This mission has been extended to encompass an open e-infrastructure that enables flexible cooperation and optimal use of all electronically available resources, paving the way towards a general-purpose European e-infrastructure. The main e-infrastructure components and services include networking, high-throughput and high-performance computing, data infrastructures, software/middleware (including authentication and authorisation infrastructures) and virtual research environments that are to be used by international VRCs. As a consequence, significant attention and funding has been put into the development of a roadmap for European research infrastructures (ESFRI[30]) and many of these research infrastructures are now becoming operational.

In November 2011, the European Commission proposed the new Common Strategic Framework Programme for research and innovation (Horizon2020) the shaping of which will continue until its start-up in 2014. Horizon2020 is not limited to just academic or public research. Its goals can only be achieved if the e-infrastructures support the entire innovation chain.

The challenges e-IRG and Horizon2020 are facing relate mainly to data-intensive sciences (d-Science) and appropriate infrastructures for data services such as management issues, sustainability issues, legal issues, access, security, interoperability, etc.; and to the coordination between different actors *('components')* in the field of e-infrastructures.

From this perspective, the work performed by MAPPER (i.e., the MAPPER services) can play an important role in supporting d-science activities and component coordination.

8 Action plan

In order to implement both the MAPPER Sustainability Plan and Standardization Roadmap, MAPPER proposes an organisational setup and several dedicated actions.

[30] http://ec.europa.eu/research/infrastructures/index_en.cfm?pg=esfri

The organisational setup defines the (structural and procedural) context of the actions to be performed. During the project's runtime the context is implicitly defined by the project's structure. After the end of the project it is proposed to create a MAPPER virtual organisation with the objective to further promote the MAPPER concepts and results (Action 12 in Table 2).

The primary sustainability actions are listed in Table 2. It should be noticed that these actions represent generic activities, which need to be translated into concrete activities. For example, Action 9 refers to several standardization bodies. It should be noted that the MAPPER beneficiary leading this Action is implicitly determined by the supporting deliverables (column (e) in the table).

The items to be sustained are listed in Appendix B.

Table 2: Actions to implement the sustainability plan

Id	Action	Action to be performed		Primary source of information (if delivered during the project)	Primary source of information (if delivered after the project)
		during the project	after the project		
(a)	(b)	(c)	(d)	(e)	(f)
1	**Provide a description of all necessary resources**	x		D4.1, D6.2, D7.3, D8.1	
2	**Provide a description of all necessary and available services and software packages**	x	x	D5.1, D6.2, D8.1	Resource and Service Providers (using a MAPPER provided template)
3	**Provide a user manual of how to get access to and use the resources and services**	x	x	D8.4	Resource and Service Providers (addenda to the MAPPER manual)
4	**Provide a support manual**	x		D6.3	
5	**Provide an SLA template**	x		D6.2	
6	**Negotiate MoUs**	x		ongoing	
7	**Present MAPPER results**	x	x	D2.2, D2.3	MAPPER final results presentation
8	**Provide MAPPER training material**	x		D2.5.1, D2.5.2	
9	**Representation in standardization bodies**	x	x	ongoing	personal participation
10	**Definition and promotion of MAPPER profile**	x		D3.3, D3.5	
11	**Publish scientific results obtained from applying MAPPER tools and services**	x	x	ongoing	
12	**Create a MAPPER VO consisting of MAPPER partners and interested stakeholders**		x		to be defined in month 36
13	**Apply for follow-on projects**		x		to be defined in month 36

9 Conclusion

This document describes the Final Sustainability Plan of the MAPPER project. Its purpose is to ensure that the results of the MAPPER project are adopted and further developed after the end of the project.

MAPPER's approach to sustainability is structured into the four layers of tasks and activities. Each layer is associated with various standardization and sustainability tasks, activities and the relevant communities. The four layers are: *science* (science communities employing MMS applications); *ICT R&D* (computer science communities developing new MMS concepts and methods); *software* (ICT developers concerned with high-level and low-level software components, tools and services); and *resources* (e-infrastructures providing the ICT resources for MMS applications).

This document represents the final version of the Sustainability Plan. It (a) provides an update of the dimensions relevant to sustainability, (b) discusses initial outcomes resulting from the implementation of the Plan, and (c) outlines future implementation activities.

The previous version of the MAPPER Sustainability Plan and other supporting documents are available through the project website at http://www.mapper-project.eu.

References

Dav05 Davidson, E. & Levin, M.: Gene regulatory networks. Proceedings of the National Academy of Sciences of the United States of America, 102(14):4935, 2005.

Fos01 Foster, I.; Kesselman, C.; Tuecke, S.: The Anatomy of the grid: Enabling Scalable Virtual Organizations; International Journal of High Performance Computing Applications, 2001, 15, 200-222.

Hoo06 Hoops, S., Sahle, S., Gauges, R., Lee, C., Pahle, J., Simus, N., Singhal, M., Xu, L., Mendes, P. & Kummer, U.: COPASI – a Complex Pathway Simulator, Bioinformatics, 22(24):3067–3074, 2006.

IRG10 e-Infrastructure Reflection Group (e-IRG): e-IRG Roadmap 2010. www.e-irg.eu/images/stories/eirg_roadmap_2010_layout_final.pdf

MAP10 Deliverable D2.2 of the MAPPER project: Final Plan for the Use and Dissemination of Foreground, 2010. www.mapper-project.eu/documents/

MAP10a Deliverable D3.1 of the MAPPER project: Report on the Policy Framework Resource Providers Need to Adopt to Support the MAPPER Project, 2010.

MAP11a Deliverable D4.1 of the MAPPER project: Review of Applications, Users, Software and e-infrastructures, 2011.

MAP11b Deliverable D6.1 of the MAPPER project: Report on the Assessment of Operational Procedures and Definition of the MAPPER Operational Model, 2011. www.mapper-project.eu/documents/

MAP11c Deliverable D6.2 of the MAPPER project: Report on Best Practices, 2011. www.mapper-project.eu/documents/

MAP11d Deliverable D3.2 of the MAPPER project: Standardization Roadmap and First Sustainability Plan, 2011. www.mapper-project.eu/documents/

MAP11e Deliverable D3.3 of the MAPPER project: MAPPER Profile, 2011. www.mapper-project.eu/documents/

MAP11f Deliverable D5.2 of the MAPPER project: MAPPER Vertical Integration Plan, 2011. www.mapper-project.eu/documents/

MAP11g Deliverable D5.1 of the MAPPER project: Report on the Inventory of Deployed Services, 2011. www.mapper-project.eu/documents/

MAP12a Deliverable D3.5 of the MAPPER project: MAPPER Test Suite, 2012. www.mapper-project.eu/documents/

Maz08 Mazzeo, M.D., Coveney, P.V. (2008) HemeLB: A high performance parallel lattice-Boltzmann code for large scale fluid flow in complex geometries. Comput Phys Commun, 178 (12) 894-914.

Men11 Dada, J.O.: Mendes, P. Multi-scale modelling and simulation in systems biology, Integrative Biology,3, 86-96, 2011

Appendix A: Abbreviations

API	Application Programming Interface
BES	Basic Execution Service
BioPAX	Biological Pathway Exchange
DANTE	Delivery of Advanced Networking Technology to Europe
DICOM	Digital Imaging and Communcations in Medicine
DMC	Distributed Multiscale Computing
DMTF	Distributed Management Task Force
DRMAA	Distributed Resource Management Application API
EGI	European grid Infrastructure
EGI-InSPIRE	EGI Integrated Sustainable Pan-European Infrastructure for Researchers in Europe
e-IRG	e-infrastructure Reflection Group
EIRO	European International Research Organisation
ESFRI	European Strategy Forum on Research Infrastructures
ExTENCI	Extending Science Through Enhanced National Cyberinfrastructure
FieldML	Field Modelling Markup Language
FTP	File Transfer Protocol
GO	Gene Ontology
gSLM	Service Delivery and Service Level Management in grid Infrastructures
HPC	High Performance Computing
ICT	Information and Communication Technology
IETF	Internet Engineering Task Force
MAPPER	Multiscale Applications on European e-infrastructures
MBC	MAPPER Multiscale Base Case
MITA	Medical Imaging and Technology Alliance
MMS	Multiscale modelling and simulation
MoU	Memorandum of Understanding
NEMA	Association of Electrical and Medical Imaging Equipment Manufacturers
NGI	National Grid Infrastructure
NREN	National Research and Education Network
OASIS	Organization for the Advancement of Structured Information Standards
OBO	Open Biological and Biomedical Ontologies
OGF	Open Grid Forum
OGSA	Open Grid Services Architecture
OMG	Object Management Group
PRACE	Partnership for Advanced Computing in Europe
QosCosGrid	Quasi-opportunistic supercomputing for complex systems in Grid environments
SAGA	Simple API for Grid Applications
SBGN	Systems Biology Graphical Notation
SBML	Systems Biology Markup Language
SLA	Service Level Agreement
SUT	System Under Test
VO	Virtual Organization
VRC	Virtual Research Community
W3C	World Wide Web Consortium
WP	Work Package
XSEDE	Extreme Science and Engineering Discovery Environment

Appendix B: List of MAPPER Components to be Sustained

Table 3: Items to be sustained

Name	Description (Function, Purpose)	Support/Sustainability	Partner(s)	Remarks
Application Hosting Environment (AHE)	Framework for providing user-friendly means of executing applications on remote resources.	AHE is under active development and maintenance in VPH-Share until 2015, and will remain available as a public download.	UCL	-
GridSpace	Web-based platform enabling researchers to execute multiscale applications in a form of virtual experiments on Grid-based resources and HPC infrastructures.	Maintaining the service, providing it as open-source distribution and source code via web page. Respond to queries. Use and further develop in future projects.	CYFRONET	-
GRN models	GRN models as SBML-formatted files.	Provide "as is" to computational biology via UU/MAPPER website or SBML repository. Respond to queries.	UU	-
HemeLB	A richly-featured lattice-Boltzmann blood flow simulator used in the multiscale cerebrovascular blood flow application.	HemeLB remains under active development at least until 2015, through several EPSRC and EU projects. The last one to end is the EPSRC project "Large Scale Lattice-Boltzmann simulation of BioColloids".	UCL	-
MAPPER Memory	Web framework enabling researchers to register information about single scale modules and the way they are connected forming multiscale applications (using XMML).	Maintain the framework and provide it as open-source via publicly available GForge Cyfronet server with open read access. Respond to queries. Use and further develop in future projects.	CYFRONET	-
MPWide	C++-based communication library for wide area message passing.	Maintain the library, and maintain a public download location of the source code, under an LGPL license. The lead developer is willing to provide informal support at least until 2016. Additionally, MPWide is incorporated into MUSCLE 2.0.	UCL	-

Name	Description (Function, Purpose)	Support/Sustainability	Partner(s)	Remarks
MultiGrain (GRN modeling toolbox)	Java library facilitating modeling/simulation of gene-regulation networks.	Maintain library and provide it as open-source via SourceForge. Respond to queries. Use and further develop in future projects.	UU	-
Multiscale Application Designer	Web framework enabling researchers to compose multiscale applications from existing single scale modules	Maintain the framework and provide it as open-source via publicly available GForge Cyfronet server with open read access. Respond to queries. Use and further develop in future projects.	CYFRONET	-
Multiscale Coupling Library and Environment (MUSCLE)	Java and C++ library, API, and execution environment implementing and running multiscale models. It runs on single computers, clusters, and supercomputers; it includes a TCP/IP communication tool for distributed computing.	Maintain software and provide it as open-source via the PSNC software repository. Respond to queries. Use and further develop in future projects.	UvA, PSNC	Under LGPLv3 license
Nanomaterials coupling framework	A collection of scripts we use for coupling the submodels in the Nanomaterials application.	As this framework is very new and under heavy development, we have not yet decided how we wish to sustain it.	UCL	-
QosCosGrid	An integrated system offering advanced job and resource management capabilities to deliver to end-users supercomputer-like performance together with various application tools and graphic interfaces.	QosCosGrid middleware stack is deployed on all PL-Grid (a polish NGI) resources which has to be maintained for the next 5 years. Moreover, the software solutions are continuously enhanced within ongoing projects, in particular PLGrid PLUS (scheduled end date: 2015). PSNC has been identified as a technology provider to support QosCosGrid in Distributed Computing Infrastructures in Europe, in particular EGI and PRACE.	PSNC	
Synthetic dynamic gene expression data sets	Data sets (SBRML) generated via artificial GRN systems.	Provide “as is” to computational biology via UU/MAPPER website. Respond to queries.	UU	-

Key/Abbreviations:

Name

Succinct name of sustained item

Description (Function, Purpose)

Describes the nature of the item. What is it (SW, components, API, service, tool, method, document, etc.) and what purpose and function it has.

Support/Sustainability

What kind of support or sustainability will be provided beyond the life time of the project, for how long and under which conditions.

Partner(s)

Acronyms of beneficiary names (from list below) or other organizations (if any) who will provide the support or sustain the item.

UvA	University of Amsterdam
UCL	University College London
UU	University of Ulster
PSNC	Poznan Supercomputing and Networking Center
Cyfronet	Cyfronet AGH
LMU	Ludwig-Maximilians-Universität München
UNIGE	Université de Genève
CTHA	Chalmers University of Technology
MPI-IPP	Max-Planck-Institut für Plasmaphysik