Deliverable D5.1

Report on the Inventory of Deployed Services

Project acronym: *MAPPER*

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

Grant agreement no.: 261507

Due-Date:	Month 24
Delivery:	Month 24
Lead Partner:	PSNC
Dissemination Level:	PU
Status:	Draft
Approved:	
Version:	2.6

DOCUMENT INFO

Date and version number	Author	Comments
10.03.2011 v1.0	Mariusz Mamoński	Initial version
12.03.2011 v1.1	Bartosz Bosak	Added missing citations
14.03.2011 v1.2	Bartosz Bosak	Summary given
15.03.2011 v1.3	Mariusz Mamoński	Fixed formatting
15.03.2011 v1.4	Bartosz Bosak, Mariusz Mamoński	QA comments addressed
30.03.2011 v1.5	Mariusz Mamoński	Deployment diagram added
31.03.2011 v 1.6	Krzysztof Kurowski	Final version (Month 6)
7.03.2012 v 2.0	Bartosz Bosak	First Update (Month 18)
28.08.2012 v 2.1	Mariusz Mamoński	New services added (SAGA BigJob and MTO)
06.09.2012 v 2.2	Bartosz Bosak, Mariusz Mamoński	Minor corrections
07.09.2012 v 2.3	Bartosz Bosak, Mariusz Mamoński	Updated “Deployed services matrix”
07.09.2012 v 2.4	Bartosz Bosak,	Minor corrections
07.09.2012 v 2.5	Bartosz Bosak	Final editing (Month 24)
14.09.2012 v 2.6	Bartosz Bosak	Addressed internal reviewers’ comments

TABLE OF CONTENTS

LIST OF TABLES AND FIGURES

Executive summary

This deliverable is a living document serving as the report on the inventory of deployed and exploited middleware services in the context of the MAPPER project, in particular advanced multi-scale applications and their requirements that can not be fulfilled by existing e-Infrastructures. This version of report is a snapshot and it lists all the relevant middleware services deployed on European e-Infrastructures at the end of month 24 of the project. The main focus of this document is to present useful functionalities of different middleware services and tools that are not a regular part of the available European e-Infrastructures: EGI and PRACE.

In principle, all the initially deployed middleware services in the MAPPER project are expected to extend capabilities provided by the existing e-Infrastructures and improve their interoperability. Nevertheless, the main goal of newly deployed middleware services and tools is to meet both specific needs and requirements of the MAPPER multi-scale applications. The list of described middleware services is not limited to services classified in project Description of Work [1] as the “fast track" components (i.e. tools identified as the minimal set of infrastructure components enabling the coupling of multi-scale applications), but it also includes tools classified as the "deep track" components (i.e. tools that realize fully automatic coupling and launching of multi-scale applications). The deliverable shortly describes the key middleware services deployed on the pre-production MAPPER infrastructure, namely: AHE, SPRUCE, QCG-Computing, QCG-Broker, SAGA BigJob, MUSCLE Transport Overlay and GridSpace. It also presents in the matrix form the availability of particular middleware services at given sites that compose MAPPER pre-production testbed.

The current version of the report extends its previous release in a number of sites where the MAPPER components were deployed. Since the first version of the document (M6) we decided not to use HARC as its development had been stopped completely. Fortunately the HARC service could be replaced with the QCG-Computing service without any losses in functionality. On the other side, we decided to support two new services: SAGA BigJob and MUSCLE Transport Overlay (MTO), both described in the further sections.

The report includes also one additional section devoted to best practices and procedures related to deployment of the MAPPER services.

Finally, this report is a complementary document to other complementary reports that are published at the same time in Service Activities, namely:

- D4.1 Review of Applications, Users, Software and e-Infrastructures, and
- D6.1 Report on the Assessment of Operational Procedures and Definition of the MAPPER Operational Model.

Update notes

The document is periodically extended by description of further deployments of particular services. During the project lifetime the document was modified in the following way:

PM 24 (version 2.3)

1. Described new services: Saga BIG Jobs and MTO.
2. Updated Mapper testbed section and Deployed services section along with the matrix presenting deployment of the selected services on sites.
3. Minor corrections.

Contributors

- PSNC: Krzysztof Kurowski, Mariusz Mamoński, Bartosz Bosak
- Cyfronet: Eryk Ciepiela
- UCL: Stefan Zasada, Derek Groen
- LMU: Ilya Saverchenko

List of Abbreviations

Item	Description
AHE	Application Hosting Environment
DEISA	Distributed European Infrastructure for Supercomputing Applications
DRMAA	Distributed Resource Management Application API
EGEE	Enabling Grids for E-sciencE
EGI	European Grid Initiative/Infrastructure
HARC	The Highly-Available Resource Co-allocator

HLRS	High Performance Computing Center Stuttgart
HPC	High-Performance Computing
JSDL	Job Submission Description Language
LFC	LCG File Catalog
NGS	National Grid Service
NW-Grid	The North West Grid
PBS	Portable Batch System
PL-Grid	Polish National Grid Initiative
QCG	QosCosGrid
SPRUCE	Special PRiority and Urgent Computing Environment
MTO	MUSCLE Transport Overlay
SAGA	Simple API for Grid Application
UNICORE	Uniform Interface to Computing Resources

Table 1. Abbreviations

Service descriptions

AHE

The Application Hosting Environment [2], AHE, developed at University College London, provides simple desktop and command line interfaces, to run applications on resources provided by national and international grids, in addition to local departmental and institutional clusters, while hiding from the user the details of the underlying middleware in use by the grid. In addition, a mobile interface for Windows Mobile based PDAs is available, and an iPhone interface is in development. The AHE is able to run applications on UNICORE [3], Globus [4] and QosCosGrid, meaning that a user can use a single AHE installation to access resources from the UK National Grid Service (NGS) [5], Polish NGI (PL-Grid) [6] and PRACE [7]. Development of European Grid Infrastructure (EGI) [8] connector for AHE is currently underway.

The AHE is designed to allow scientists to quickly and easily run unmodified, legacy applications on grid resources, manage the transfer of files to and from the grid resource and monitor the status of the application. The philosophy of the AHE is based on the fact that very often a group of researchers will all want to access the same application, but not all of them will possess the skill or inclination to install the application on remote grid resources. In the AHE, an expert user installs the application and configures the AHE server, so that all participating users can share the same application.

SAGA BigJob

The SAGA BigJob framework [9] is an abstraction over pilot jobs (i.e. a large containers that can be used for submission of a set of smaller jobs). It was developed in Center for Computation & Technology of Louisiana State University. The main advantage of SAGA BigJob, especially in context of the MAPPER project, is that it supports MPI jobs by default and also offers a wide portfolio of supported back-end systems.

The SAGA BigJob framework functionality may be used by QCG-Broker [14] and the AHE service to co-allocate Tightly Coupled Multiscale Applications on systems that do not support advance reservations and improve scheduling of Loosely Coupled Multiscale workflow jobs.

SPRUCE

The traditional high performance computing batch queue model does not allow or makes not easy for simulations to be prioritized by their urgency. Typically a grid will provide general purpose resources to a wide range of different users. If these resources are to be used by clinicians in support of their clinical practice, especially in support of emergency medical intervention planning, then some way is needed of prioritizing clinical simulations above the normal workload on a computational resource. SPRUCE, A System for Supporting Urgent High-Performance Computing [10], developed at Argonne National Labs, USA, is a tool, which allows this to happen. Clinicians and other users with simulations that are considered an emergency are issued with SPRUCE tokens, which allow them to submit emergency jobs to a machine. The SPRUCE middleware takes care of running the job in a high priority mode, pre-empting the work that is already running on the machine.

GridSpace

GridSpace [10] is a novel virtual laboratory framework enabling researchers to conduct virtual experiments including running of multi-scale applications on Grid-based resources and other HPC infrastructures. The current generation of GridSpace - GridSpace2 - facilitates the exploratory development of experiments by means of scripts, which can be expressed in a number of popular languages, including Ruby, Python and Perl. The framework supplies a repository of gems enabling scripts to interface low-level resources such as Portable Batch System (PBS) [11] queues, EGEE computing elements, LCG File Catalog (LFC) [12] directories and other types of Grid resources. Moreover, GridSpace2 provides a Web 2.0-based Experiment Workbench supporting joint development and execution of virtual experiments by groups of collaborating scientists.

MUSCLE Transport Overlay

MUSCLE Transport Overlay (MTO) [13] is an userspace daemon that makes a cross-cluster execution of parallel MUSCLE application feasible. The MTO has to be deployed at an interactive node, or any other node that is accessible from both external hosts and worker nodes, of all clusters involved in a tightly coupled multiscale simulation. Every MTO listens on a separate address for external and internal requests. The external port must be either accessible from all the other interactive nodes or the MTO must be able to connect to the external ports of all the others MTO (i.e. uni-directional connection is needed between every pair of MTOs).

QosCosGrid

QosCosGrid [14] was designed as a multilayered architecture being capable of dealing with computationally intensive large-scale, complex and parallel simulations that are often impossible to run within one computing cluster. The QosCosGrid middleware enables computing resources (at the processor core level) from different administrative domains to be virtually welded via Internet into a single powerful computing resource. QosCosGrid delivers a ready-to-use stack of grid middleware software tightly integrated with commonly used programming and execution environments for large-scale parallel simulations, such as OpenMPI [15] or ProActive [16]. Recently the QosCosGrid stack was integrated with the MUSCLE coupling library [17], a framework that is widely used by applications in the MAPPER project. Supporting a wide range of development frameworks as well as programming models relevant for multi-scale application developers, QosCosGrid gives the ability to work across heterogeneous computing sites hiding the complexity of underlying e-Infrastructures by simplifying many complex deployment and access procedures. QosCosGrid services extend the functionality provided by the gLite and Unicore infrastructures offering advance reservation capabilities needed to co-allocate various types of resources required by many of the multi-scale applications.

QCG-Computing (QCG BES/AR)

The QCG-Computing [14] service (known also as the Smoa Computing) is an open architecture implementation of SOAP Web Service for multi-user access and policy-based job control routines by various Distributed Resource Management systems. It uses Distributed Resource Management Application API (DRMAA) [18] to communicate with the underlying DRM systems. QCG-Computing has been designed and implemented in the way to support different plugins and modules for external communication. Consequently, it can be used and integrated with various authentication, authorization and accounting infrastructures and other external services. QCG-Computing service is compliant with the OGF HPC Basic Profile [19] specification, which serves as a profile over the Job Submission Description Language (JSDL) [20] and OGSA® Basic Execution Service [21] Open Grid Forum standards. In addition, it offers remote interface for Advance Reservations management, and support for basic file transfer mechanisms. The service was successfully tested with the following Distributed Resources Management systems: Sun Grid Engine (SGE) [22], Platform LSF [23], Torque/PBSPro [24], PBS Pro, Condor [25], Apple XGrid [26] and Simple Linux Utility for Resource Management (SLURM) [27]. The Advance Reservations capabilities were exposed for SGE, LSF and Maui [28] (a scheduler that is typically used in conjunction with Torque) systems.

QCG-Broker

The QCG-Broker (previously aka Grid Resource Management System - GRMS) [14] is an open source meta-scheduling system, which allows developers to build and deploy resource management systems for large scale distributed computing infrastructures. The QCG-Broker, based on dynamic resource selection, mapping and advanced scheduling methodology, combined with feedback control architecture, deals with dynamic Grid environment and resource management challenges. It is capable of load balancing of jobs among clusters and co-allocating of resources. The main goal of the QCG-Broker is to manage the whole process of remote job submission to various batch queuing systems. It has been designed as an independent core component for resource management processes which can take advantage of various low-level core and grid services responsible for execution of jobs and reservation of resources on cluster machines. The QCG-Broker allows to co-allocate resources belonging to different e-Infrastructures and execute cross-cluster and cross-infrastructure multi-scale applications.

Deployment architecture

The Figure 1 extends a bit the bottom line of the initial overall architecture of the MAPPER project called: Middleware building blocks (see Figure 3 in D4.1). It shows how the new middleware service may potentially fit into the existing e-Infrastructures as an added entity to key middleware services available in EGI and PRACE. All the MAPPER components are marked green while the traditional services are in violet. There is a clear distinction between the low level components (e.g. QCG-Computing) that must be installed locally at the resource providers side and are integrated with underlying resource management systems and the high-level one (e.g. AHE), which can be deployed on third party resources.

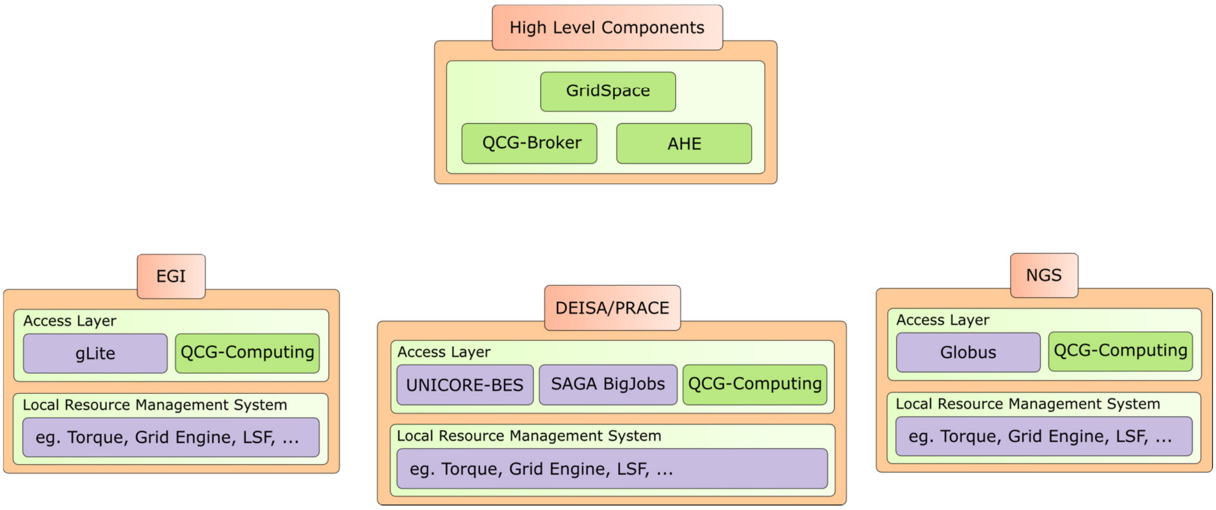

Figure 1 MAPPER deployment overview

The MAPPER testbed

The current MAPPER testbed is composed of resources and services provided by the following infrastructures:

- Local Campus Resources – hosted by University College London (UCL),
- Pl-Grid NGI (EGI) – all sites: CYFRONET, PSNC, WCSS, TASK and ICM,
- NGI_DE (EGI) – represented by LRZ,
- PRACE – represented by SARA, University of Edinburg, HLRS and PSNC.

Deployed services

The table below gives a detailed view on the current deployment of the aformentioned services. The SPRUCE tool is not mentioned in the table. This is because it is not currently deployed on any site of the MAPPER testbed. However it is installed on TeraGrid and LONI e-Infrastructures in San Diego Supercomputing Center (SDSC), in National Center for Supercomputing Applications (NCSA), in Univerirsity of Chicaga and Argone National Laboratory (UC/ANL), in Texas Advanced Computing Center (TACC) and Louisiana Tech at Ruston.

Infrastructure	Institution	System Name	GridSpace	AHE[1]	QCG Broker[2]	QCG Computing	SAGA BigJob	Muscle Transport Overlay
Campus Resource	UCL	Mavrino	Yes	Yes	Yes	Yes	No	Yes
Campus Resource	UCL	Oppenheimer	Yes	Yes	Yes	Yes	No	No
PL-Grid (EGI)	Cyfronet	Zeus	Yes	Yes	Yes	Yes	No	Yes
PL-Grid (EGI)	PSNC	Reef	Yes	Yes	Yes	Yes	No	Yes
PL-Grid (EGI)	PSNC	Inula	Yes	Yes	Yes	Yes	No	Yes
PL-Grid (EGI)	WCSS	Nova	Yes	Yes	Yes	Yes	No	Yes
PL-Grid (EGI)	TASK	Galera+	Yes	Yes	Yes	Yes	No	Yes
PL-Grid (EGI)	ICM	Hydra	Yes	Yes	Yes	Yes	No	Yes
NGI_DE	LRZ	Linux Cluster	Yes	Yes	Yes	Yes	No	No
PRACE	SARA	Huygens	Yes	Yes	Yes	No	Yes	Yes
PRACE	University of Edinburg	HECToR	Yes	Yes	Yes	No	Yes	No
PRACE	HLRS	HERMIT	Yes	Yes	Yes	No	No	No
PRACE	PSNC	Cane	Yes	Yes	Yes	Yes	No	Yes

Table 2. Deployed services matrix

Best Practices and Installation Instructions

Based on the experiences gained from deployments of the MAPPER services that have been conducted so far, a set of instructions and best practices were collected and made publicly available [29]. The aim of the prepared materials is to help administrators setup their computing environments for installation and maintenance of services that enable distributed

[1] AHE is deployed in UCL but is able to access all sites where the QCG-Computing/UNICORE (OGSA-BES interface) service is deployed.

[2] QCG-Broker is able to submit jobs to all sites where the QCG-Computing/UNICORE service is deployed.

multi-scale computations. The detailed descriptions, related particularly to the QosCosGrid components, alongside with the comprehensive examples, are available on the QosCosGrid webpage[3].

Summary

All the new middleware services and tools listed in this report were selected due to useful features they provide for multi-scale application use-cases, in particular for tightly-coupled and loosely-coupled multi-scale scenarios considered in the MAPPER project. As it was described in the section 7 of this deliverable, all the services initatily classified as “fast track” or “deep track” components, except SPRUCE (available mostly in the TeraGrid e-Infrastructure in the United States), have been successfuly deployed on selected sites across Europe. All resource providers involved in the project supported the current deployment phase. The addtional deployments of the MAPPER services are planed, in particular on next EGI (particulary NGI_DE, NGI_NL, NGI_UK) and PRACE sites. In order to coordinate those efforts the MAPPER-EGI-PRACE taskforce was established[4]. A first succes of the group was signing of memorandum of understanding between the EGI-InSPIRE and MAPPER projects [30]. The information about these deployments will be included in the future versions of the report. The experiences of existing deployments allowed us to create a set of best practices and installation instructions for administrators who will install the new MAPPER components in the future.

References

1 “Multiscale Applications on European e-Infrastructures” (Mapper) Project – Annex I, “Description of Work”.

2 Zasada, S. J., & Coveney, P. V. (2009). Virtualizing access to scientific applications with the Application Hosting Environment (Vol. 180). Computer Physics Communications.

3 Uniform Interface to Computing Resources, http://www.unicore.eu/.

4 The Globus Alliance, http://www.globus.org/.

5 National Grid Service, http://www.ngs.ac.uk/.

[3] http://www.qoscosgrid.org/trac/qcg/wiki/installation

[4] https://wiki.egi.eu/wiki/MAPPER-PRACE-EGI_Task_Force_(MTF)

6 Polish National Grid Initiative, http://www.plgrid.pl

7 Partnership for Advanced Computing in Europe, http://www.prace-project.eu/

8 European Grid Infrastructure, http://www.egi.eu/.

9 SAGA BigJobs, https://github.com/saga-project/BigJob/wiki

10 Ciepiela, E., Harezlak, D., Kocot, J., Bartynski, T., Kasztelnik, M., Nowakowski, P., et al. Programming in the Virtual Laboratory. Proceedings of the International Multiconference on Computer Science and Information Technology, (pp. 621-628).

11 PBS Professional, http://www.pbsworks.com/Product.aspx?id=1

12 https://twiki.cern.ch/twiki/bin/view/EGEE/GliteLFC/

13 Borgdorff J., Bona-Casasa C., Mamonski M., Kurowski K., Piontek T., Bosak B., Rycerz K., Ciepiela E., Gubala T., Harezlak D., Bubaka M., Lorenz E., Hoekstraa A.: A distributed multiscale computation of a tightly coupled model using the Multiscale Modeling Language. Simulation of Multiphysics Multiscale Systems, 9th International Workshop (to be published)

14 Kurowski, K., Piontek, T., Kopta, P., Mamoński, M., & Bosak, B. (2010). Parallel Large Scale Simulations in the PL-Grid Environment. Computational Methods in Science and Technology , 47-56.

15 Edgar Gabriel, Graham E. Fagg, George Bosilca, Thara Angskun, Jack J. Dongarra, Jeffrey M. Squyres, Vishal Sahay, Prabhanjan Kambadur, Brian Barrett and Andrew Lumsdaine, et al. Open MPI: Goals, Concept, and Design of a Next Generation MPI Implementation. Recent Advances in Parallel Virtual Machine and Message Passing Interface. pp. 353-377 (2004).

16 Caromel, D., Delbe, C., di Costanzo, A. and Leyton, M.: ProActive: an Integrated Platform for Programming and Running Applications on Grids and P2P systems. Computational Methods in Science and Technology, vol. 12, no. 1, pp. 69-77 (2006).

17 Multiscale Coupling Library and Environment, http://muscle.berlios.de/.

18 Peter Troeger, Hrabri Rajic, Andreas Haas. Standardization of an API for Distributed Resource Management Systems. Proceedings of the 7th IEEE International Symposium on Cluster Computing and the Grid (CCGrid'07), Rio De Janeiro, Brazil, May 2007.

19 GFD 114-HPC Basic Profile, Version 1.0, http://www.ogf.org/documents/GFD.114.pdf.

20 GFD 56 - Job Submission Description Language (JSDL) Specification, Version 1.0 - http://www.gridforum.org/documents/GFD.56.pdf.

21 GFD 108 - OGSA® Basic Execution Service Version 1.0, http://www.ogf.org/documents/GFD.108.pdf.

22 Oracle Grid Engine, http://www.oracle.com/us/products/tools/oracle-grid-engine-075549.html.

23 Platform Load Sharing Facility, http://www.platform.com/workload-management/high-performance-computing.

24 Torque Resource Manager, http://www.clusterresources.com/pages/products/torque-resource-manager.php.

25 Condor High-Throughput Computing System, http://www.cs.wisc.edu/condor/.

26 Apple Xgrid, www.apple.com/server/macosx/technology/xgrid.html.

27 Simple Linux Utility for Resource Management (SLURM), https://computing.llnl.gov/linux/slurm/.

28 Maui Scheduler, http://www.clusterresources.com/products/maui/.

29 MAPPER Deliverable D6.2 Report on Best Practices, http://www.mapper-project.eu/documents/10155/23478/D6.2-Report_on_Best_Practices.pdf

---

30 Memorandum of Understanding between EGI-InSPIRE and MAPPER https://documents.egi.eu/public/RetrieveFile?docid=493&version=8&filename=EGI-InSPIRE-MOU-MAPPER-FINAL.pdf