
# 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 6 |
| **Delivery:** | Month 6 |
| **Lead Partner:** | PSNC |
| **Dissemination Level:** | PU |
| **Status:** | Final |
| **Approved:** | Quality Assurance Board, Project Steering Group |
| **Version:** | 1.6 |

## 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,<br>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 |

## TABLE OF CONTENTS



## LIST OF TABLES AND FIGURES

# 1 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. The report is a snapshot and it lists all the relevant middleware services deployed on European e-Infrastructures at the end of month 6 of the project. The main focus of this document is to present useful functionalities of different middleware services and tools that are not 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, despite of the early stage of the project, 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 what is the target scenario). The deliverable shortly describes the key middleware services deployed on the pre-production MAPPER infrastructure. The report presents in the matrix form the availability of particular middleware services at given sites included in MAPPER. The current version of this report refers to the following middleware services: AHE, HARC, SPRUCE, QCG-BES/AR, QCG-Broker and GridSpace. Future versions of this deliverable will extend this list both in terms of the number of middleware services as well as their possible integration and deployment with middleware services provided in EGI and PRACE. Additionally, we will update the number of computing resource providers where the respective middleware services and tools are deployed and supported.

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.

# 2 Contributors

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

# 3 List of Abbreviations

| Item | Description |
|---|---|
| AHE | Application Hosting Environment |
| DEISA | Distributed European Infrastructure for Supercomputing Applications |
| EGEE | Enabling Grids for E-sciencE |
| EGI | European Grid Initiative/Infrastructure |
| gLITE-WMS | gLite Workload Management System |
| HARC | The Highly-Available Resource Co-allocator |
| 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 |
| UNICORE | Uniform Interface to Computing Resources |

**Table 1. Abbreviations**

# 4 Service descriptions

## 4.1 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 both UNICORE [3] and Globus [4] grids, meaning that a user can use a single AHE installation to access resources from the UK National Grid Service (NGS) [5] and the Distributed European Infrastructure for Supercomputing Applications (DEISA) [6] for example. Development of an European Grid Infrastructure (EGI) [7] 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.

## 4.2 HARC

HARC, the Highly-Available Resource Co-allocator [8] developed at the Louisiana State University is an extensible, open-source system for creating and managing resource reservations. HARC uses Paxos Commit [9], a transaction commit protocol, to reserve multiple resources in a single, indivisible step. The most common use of HARC is to make advance reservations on multiple supercomputers. This functionality will be used by QCG-Broker [13] to co-allocate NGS resources for running cross-cluster multi-scale applications by the AHE service.

## 4.3 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.

## 4.4 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.

## 4.5 QosCosGrid

QosCosGrid [13] 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 [14] or ProActive [15]. Recently the QosCosGrid stack was integrated with the MUSCLE coupling library [16], a framework which is widely used by applications in the MAPPER project. Supporting a wide range of development frameworks as well as programming models relevant for multiscale 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.

### 4.5.1 QCG-BES/AR

The QCG-BES/AR [13] 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) [17] to communicate with the underlying DRM systems. QCG-BES/AR 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-BES/AR service is compliant with the OGF HPC Basic Profile [18] specification, which serves as a profile over the Job Submission Description Language (JSDL) [19] and OGSA® Basic Execution Service [20] 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) [21], Platform LSF [22], Torque/PBSPro [23], PBS Pro, Condor [24], Apple XGrid [25] and Simple Linux Utility for Resource Management (SLURM) [26]. The Advance Reservations capabilities were exposed for SGE, LSF and Maui [27] (a scheduler that is typically used in conjunction with Torque) systems.

### 4.5.2 QCG-Broker

The Grid Resource Management System (GRMS aka QCG-Broker [13]) 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.

# 5 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 existing services are in violet. There is a clear distinction between the low level components (e.g. QCG-AR) 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.

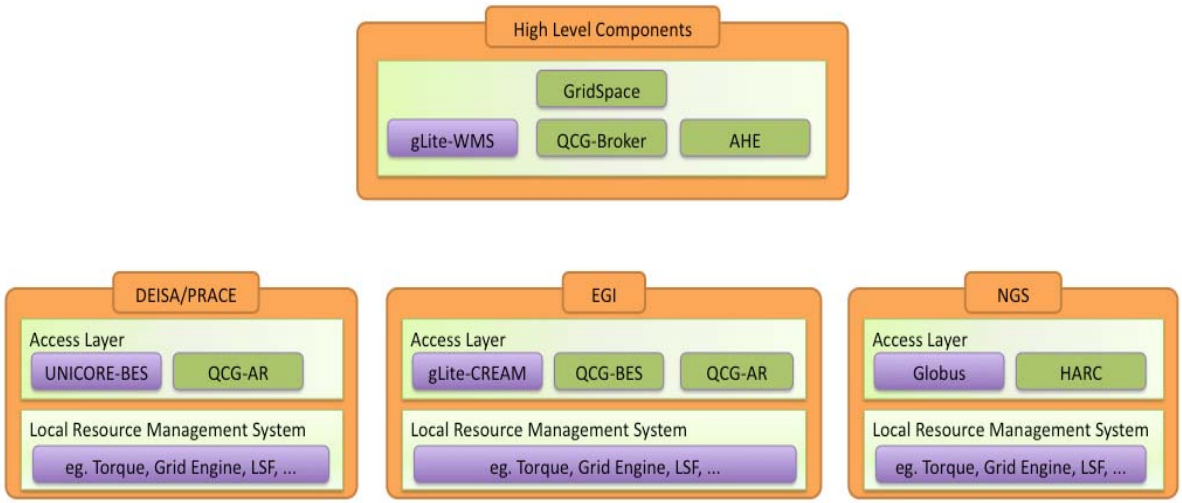


**Figure 1 MAPPER deployment overview**

# 6 The MAPPER testbed

The initial MAPPER testbed is mostly composed of resources and services provided by the MAPPER consortium:

- NW-Grid and NGS – represented by UCL
- Pl-Grid NGI – represented in MAPPER by CYFRONET and PSNC.

# 7 Deployed services

The table below gives a detailed view on the 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 intial 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 | AHE[1] | HARC | GridSpace | QCG-BES/AR | QCG-Broker[2] |
|---|---|---|---|---|---|---|
| NW-Grid | University of Manchester | No | No | No | No | No |
| NW-Grid | University of Lancaster | No | No | No | No | No |
| NW-Grid | University of Liverpool | No | No | No | No | No |
| NW-Grid | Daresbury Laboratory | No | No | No | No | No |
| NGS | University of Leeds | No | Yes | No | No | No |
| NGS | University of Manchester | No | Yes | No | No | No |
| NGS | University of Oxford | No | Yes | No | No | No |
| NGS | Rutherford Appleton Laboratory | No | No | No | No | No |
| NGS | University of Edinburgh | No | Yes | No | No | No |
| NGS | UCL | Yes | No | No | No | No |
| EGI (PL-Grid) | Cyfronet | No | No | Yes | Yes | No |
| EGI (PL-Grid) | PSNC | No | No | No | Yes | Yes |

**Table 2. Deployed services matrix**

[1] AHE is deployed in UCL but is able to submit jobs to all NGS machines.

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

# 8 Summary

All the new middleware services and tools listed in this report were selected during the first six months due to useful features they provide for multi-scale application use-cases, in particular for tightly coupled multi-scale scenarios considered in the MAPPER project. As it was described in the previous section, 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 initatial deployment phase. The next instalation and deployment procedures are currently under discussions within the MAPPER consortium as well as with external communities, in particular with technical and operational represenatives of EGI and PRACE. Thus, the next report will provide more information about deployed MAPPER services, and their availability on prodution EGI and PRACE sites in Europe. We will also include in the next report best practices and descriptions of reference installations of MAPPER services to help administrators to extend their existing setups and support advanced multi-scale application scenarios.

# 9 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/.

6 Distributed European Infrastructure for Supercomputing Applications, http://www.deisa.eu/.

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

8 MacLaren, J. (2007). HARC: The Highly-Available Resource Co-allocator. Proceedings of GADA'07 (OTM Conferences 2007, Part II) (pp. 1385-1402). Springer-Verlag.

9 J. Gray and L. Lamport. Consensus on transaction commit. ACM TODS, 31(1):130–160, March 2006.

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 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.

14 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).

15 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).

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

17 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.

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

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

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

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

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

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

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

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

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

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