



omii europe

open middleware infrastructure institute



EU project: RI031844-OMII-Europe

Project no: **RI031844-OMII-Europe**

Project acronym: **OMII-Europe**

Project title: **Open Middleware Infrastructure Institute for Europe**

Instrument: **Integrated Infrastructure Initiative**

Thematic Priority: **Communication network development**

Deliverable JRA4.1 First report on Benchmarking Activity

Due date of deliverable: Month 12
Actual submission date:

Start date of project: **1 May 2006**

Duration: **2 years**

Royal Institute of Technology (KTH)

Revision [1.0]

Project co-funded by the European Commission within the Sixth Framework Programme (2002-2006)		
Dissemination Level		
PU	Public	
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

Document Control Sheet

Document	Title: First Report on Benchmarking Activity	
	ID: D:JRA4.1	
	Version: 1.0	Status: Final
	Available at: http://omii-europe.org	
	Software Tool: Microsoft Word 2004 for Mac	
	File(s): DJRA4.1.doc, DJRA4.1.pdf	
Authorship	Written by:	Per Alexius (KTH) B. Maryam Elahi (KTH) Giuseppe Fiorentino (INFN) Fredrik Hedman (KTH) Philip Mucci (KTH) Gilbert Netzer (KTH)
	Contributors:	KTH, INFN
	Reviewed by:	<u>Tim Parkinson (SOTON)</u> Sergio Andreozzi (INFN)
	Approved by:	Technical Committee

Document Status Sheet

Version	Date	Status	Comments
0.51	27 April 2007	Draft	Version to be sent to internal reviewers.
0.52	3 May 2007	Draft	First internal review modifications.
0.53	16 May 2007	Draft	Added Partners and Efforts section.
0.54	17 May 2007	Draft	Refined and updated Partners Effort section.
0.55	23 May 2007	Draft	Included comments on Summary section.
0.60	25 May 2007	Reviewable	Included comments on all sections.
1.0	2007-06-05	Final	Many review comments included.

Executive Summary

The JRA4 activity is targeted at performance evaluation of Grid middlewares. The goals of the activity are to provide tools for assessing the performance of Grid components, to provide comparative performance data across different middlewares and to develop tutorial material for the tools to extract such data.

This document is the first deliverable for JRA4 and the only one for month twelve. It describes what has been performed during the first year as well as the plans for the coming year.

During the first year the initial version of a set of tools for extracting performance data for different middlewares has been developed. An extensive literature survey has been performed in order to create a good overview of the current status in the field of Grid benchmarking. An evaluation infrastructure has been acquired and set up and middlewares installed in order to meet the comparability requirement. The performance assessment tools have been used on this testbed to produce initial comparative performance data across gLite, Globus and UNICORE, mainly focused on job submission. A first version of tutorial material for the tools has also been developed. The plans for the second year of JRA4 includes the development of software libraries for instrumenting Grid components in order to extract performance data with higher level of detail as well as the infrastructure for storing such data and making it available to users, developers and administrators in suitable formats. The tools produced during the first year will also be improved and extended to cover more use cases.

Table of Contents

1	Introduction.....	6
2	Partners and Efforts.....	6
3	Progress.....	7
3.1	Phase 1—Progress During Year 1	7
3.2	Phase 2—Second Year Plans	8
4	Summary	10
5	References.....	11

1 Introduction

Open Middleware Infrastructure Institute for Europe aims at bringing together the best Grid technologies from Europe and elsewhere and making them available to scientists across the European Research Area, providing them with ease of use, support and quality assurance. The project facilitates the development of an open Grid infrastructure by porting and re-engineering components to adopt public standards and extend interoperation between the chosen Grid distributions.

To ensure the highest possible throughput to the users and to provide the OMII engineers with the information needed to focus their efforts, a thorough performance assessment study of the Grid middlewares and their components must be performed. Thus, an entirely new suite of performance tools specific to the Grid is required. The objective of the forth joint research activity (JRA4), benchmarking aims to provide a means for a comparative performance assessment of all the Grid components used in the middlewares of OMII-Europe. This includes the implementation of standard-based performance assessment tools that is usable by Grid developers, Grid operators and Grid users, the publication of results from using these tools across the middlewares in OMII Europe, as well as tutorial material for the tools.

This document is the first deliverable of JRA4. The purpose of this document is to define the activity's approach to Grid middleware performance analysis and benchmarking. Furthermore, it provides a record of the first twelve months of the activity and outlines the plans for the coming year.

2 Partners and Efforts

This activity is led by KTH and will involve 5.95 staff years of effort over two years, INFN 2.0 (1.0 funded, 1.0 unfunded), KTH 2.85 (2.85 funded), BU 0.5 (0.5 unfunded) and CNIC 0.5 (0.5 unfunded). In October 2006 CNIC officially withdrew from JRA4. Table 1 shows the efforts from participating partners during the first project year in staff months and the fraction of effort that should have been spent during the first year. Due to a combination of initial staff shortage and employees leaving the project it is clear that the effective staff efforts will need to increase during the coming 12 months. As is clear from the table, there are currently two active partners: KTH and INFN. Both these partners will increase the effort within this task and we will also try to activate the BU partner. Furthermore, we intend to also include a CROWN Grid expert that SOTON has offered to make available to the task. We are confident that this increased effort will deliver according to the refined and updated plan presented in this report.

Org. Efforts (F/U)	KTH	INFN		BU	CNIC
	1.425	0.50	0.50	0.25	0.25
May-06	0.000	0.00	0.00	0.00	0.00
Jun-06	0.000	0.00	0.00	0.00	0.00
Jul-06	0.000	0.50	0.00	0.00	0.00

Aug-06	0.000	0.50	0.00	0.00	0.00
Sep-06	0.400	0.50	0.00	0.00	0.00
Oct-06	1.425	0.50	0.00	0.00	
Nov-06	1.425	0.50	0.00	0.00	
Dec-06	1.425	0.50	0.00	0.00	
Jan-07	1.425	0.50	0.00	0.00	
Feb-07	1.425	0.50	0.00	0.00	
Mar-07	1.425	0.50	0.00	0.00	
Apr-07	1.925	0.50	0.00	0.00	
Sum FTE	10.875	5.000	0.00	0.00	0.00

Table 1 – Overview of efforts during the first project year, in staff months. (F/U) is funded and unfunded respectively. The fraction represents the fraction of the effort for the first 12 months that has actually been spent.

3 Progress

The goals of JRA4 is to address:

- The integration of performance assessment tools into the OMII-Europe repository;
- An assessment of the performance of re-engineered and ported components available in the OMII- Europe Repository;
- The development of tutorial material for the above tools. This will link into the training networking activity.

The above points imply that we need to provide performance assessment tools that makes it possible to produce comparative performance data across different middlewares. As an “honest broker” we also need to produce such data and make it publicly available. We have chosen to approach this task in two phases. The following sub-sections briefly describe the two phases.

3.1 Phase 1—Progress During Year 1

The first phase consisted of a significant amount of information gathering. An extensive literature survey [2] was performed in order to create an overview of current work in the field. This work will continue during the second phase, including new information as it becomes available. We found that the area of Grid middleware performance analysis was not very well described in the published literature. While some work existed, it was very narrow in scope and not comparative in nature. A basic requirement for any experiment is to have a controlled environment. In this case, this meant that performance comparisons must be run on exactly the same hardware and operating system release. Thus in phase 1, JRA4 acquired and set up a testbed to be used for the performance measurements. The process has been an important one, allowing the JRA4 developers to become familiar with the installation and configuration of each of the middlewares and understanding how they interact with the system. Theoretically, the installation of these Grid systems should be a simple task, but we found the opposite true and much help was obtained by consulting the respective middleware communities. The test bed currently consists of four dual processor nodes

running Scientific Linux 3.0.8. On these nodes, gLite 3.0.1, Globus 4.0.4 and UNICORE 5 have been installed and installation of CROWN Grid is in progress.

Phase 1 has also involved creation of the first version of a set of tools that can be used to obtain and visualize performance data. Two important design goals for these tools have been to make them non-invasive and independent of the underlying Grid middleware. The tools consist of scripts for submitting jobs to the different middlewares, recording the necessary timing information and for processing the output into graphs and charts. Currently the results produced by these tools are being validated. Once this process is complete in M14, the tools will be integrated in the OMII-Europe repository.

On all Grid systems, jobs are submitted, executed and the output is returned. This use case was chosen as a starting point for the measurements performed in the scope of the first phase of JRA4. Information extracted from this includes how much overhead the Grid imposes on a submitted job and how this overhead changes when the number of submitted jobs varies. The idea behind these tools is to look at a Grid installation as a black box into which jobs are submitted and from which results are returned. Recording timestamps for when jobs are submitted, started, terminated and for when the output is returned, makes it possible to measure the overhead imposed on jobs by a middleware prior to and after execution. This information gives an initial view of how different middlewares perform in comparison to each other in this use case and may also give a starting point for where to focus on further, more detailed measurements. Up until now, the jobs that have been submitted are ‘no-op’ jobs, i.e. jobs with a negligible execution time. A natural continuation would be to extend this to include also other types of jobs in order to stress other parts of the systems. The tools have been used together with the middlewares installed on the JRA4 testbed in order to perform initial comparative performance measurements. Results from these measurements can be found in [1] and have already resulted in code changes in Globus through a very positive collaboration. Over time we hope to collaborate with other Grid middleware teams in the same positive way. Phase 1 has also included the creation of tutorial material for the tools in order to facilitate for interested parties to use them.

3.2 Phase 2—Second Year Plans

Phase 1 will continue during the second year of the project. CROWN Grid will be installed and the tools extended to handle it. Furthermore the tools will also be validated. We will also introduce additional jobs beyond the “no-op” job currently in use, jobs requiring I/O being of particular interest.

The plan for phase 2 has gradually evolved during phase 1, the end goal being to provide much more detailed performance information about each of the Grid components as well as comparative performance data across the different middlewares. It should be noted that there is an extensive amount of software to deal with, each of which differs in interface, semantics, complexity and available documentation. We propose a system that, through the use of a multi-language instrumentation API, records timestamps whenever a job or components of a job are serviced, manipulated or otherwise change state. (*i.e.* authentication, file-staging, etc.) This data will then be stored in a performance database (PDB) for access by other parts of the infrastructure. In order to facilitate delivery of this information to the user in a standardized way, we will use the Usage Record (UR), a report that contains meta-data about a job's execution. The UR specification supports extensions, which we hope to utilize to include performance data gathered during the lifetime of a job, which we call a Performance Record (PR). The format of a PR is yet to be defined

by JRA4, but nature of the extension mechanism of UR's makes it possible for components not interested in PR's to ignore them without requiring any code changes. The PR's will be generated from data stored in the PDB. The PR's will then be attached to the UR and after some time, be removed from the PDB. Tools to extract the performance data from UR's and produce readable output for a variety of audiences will be developed.

Complete instrumentation of all the components of the middlewares is not possible within the scope of JRA4. As such, we will focus on the components that contribute significantly to an applications runtime with guidance from the original developers of the middleware under study.

Phase 2 will also continue the evaluation of the existing toolset to determine potential usefulness, including ServMark. Lastly, we will investigate the possibility of integrating performance regression tests for individual Grid components into the OMII repository and the ETICS build and test system.

As a continuation of phase 1, the second year will include the installation of CROWN Grid on the testbed and extension of the tools in order to be able to evaluate it. The tools will also be further developed so that Condor-G can be used as submission mechanism for Globus, and so that they can handle use cases with a focus other than job submission. Additional performance data will be gathered using the extended tools. These tasks will be accomplished by the end of M14.

The first task of phase 2 will be the definition of the format of PR's and how UR's should be extended to contain them. It will be finished at the end of M15. The development of the instrumentation libraries and the PDB will then be the next task, which will be done by M20. In parallel with this, the evaluation of ServMark will be performed as well as an investigation as to how the performance assessment tools might be into the ETICS build and test system. These tasks will be finished by M16 and M18 respectively. An initial implementation of the tools for extracting and visualizing the performance data will be performed in parallel with the task of identifying Grid components as candidates for instrumentation. This activity will be completed by M23. M24 will be spent documenting the software and developing further training and tutorial materials. Table 2 summarizes the target milestone dates for each of the different tasks.

Task	Finished by the end of
Phase 1	
Extension of tools to handle CROWN Grid, Condor-G and more job use cases. Validation of tools.	M14
Phase 2	
Define format of PR and how UR's should be extended.	M15
ServMark evaluation	M16
Investigate ETICS performance test integration.	M18
Develop instrumentation libraries and PDB	M20
Investigate which Grid components that could be suitable to start instrumenting and instrument them.	M23
Develop tools for extracting performance information from PR's in UR's and for visualizing it.	M23
Documentation and tutorial material	M24

Table 2 - JRA4 plan for the second project year

4 Summary

The main goals of JRA4 are to provide comparative performance data across different middlewares under study in OMII, to provide tools aiding in the generation of such data, and to provide training and tutorial materials on how to use these tools and interpret the data they produce. In regards to the scope of JRA4, we claim that the measurement of highly platform-centric performance metrics (i.e. cache misses, pipeline stalls) is largely contrary to the overall goal of OMII, to provide a ubiquitous Grid-computing infrastructure. Instead, we direct our focus towards optimizing overall throughput through evaluating Grid component performance. We believe that the reduction of overhead of each Grid component imposes on a submitted job is the key to achieving better overall throughput and per-job efficiency. By providing the means to measure performance bottlenecks in the OMII middlewares, we aim to facilitate the integration of performance evaluation and optimization as an integral part of the OMII development cycle.

JRA4 consists of two phases. The bulk of the first year's effort has been spent on phase 1 and the project is now transitioning to phase 2.

Phase 1 has consisted of a significant amount of background research, strategic planning and infrastructure development. The results of this are an extensive literature survey [2], the JRA4 Whitepaper [1] and the setup of a common testbed for the evaluation of gLite 3.0.1, Globus 4.0.4 and UNICORE 5. Setting up a formal testbed was the only way to achieve scientifically valid (and reproducible) results. To our knowledge, this is the first formal Grid testbed, comprising three different Grid middlewares, ever established. Phase 1 also resulted in the development of an initial version of a toolset for extracting some simple performance data from the three middlewares. The tools were then used to perform initial performance measurements regarding the overhead of job submission and cleanup, essentially treating the Grid system as a 'block box'. Tutorial material on these tools has also been produced.

The plan for phase 2 has evolved out of the information and experience gathered in phase 1. It includes the development of an API's for the instrumentation of Grid components as well as the gathering and reporting of the performance data. Through active instrumentation, we will be able to provide performance data with a significantly higher level of detail than is achieved by phase 1. Using this API, information on the performance of each component (and it's own internal phases of computation) will be recorded during the lifetime of specifically designed test jobs. When the job is done, the recorded data will be made available as an extension to the Usage Record of the job, which is created by default. Tools will also be provided for extraction of specific and relevant performance data from the Usage Records (for end-users) as well as the underlying database (for Grid developers). In regards to the latter, we will investigate the possibility of integrating component performance tests directly into the OMII code base and the ETICS build/test system. Phase 2 will also continue the information gathering effort as well as the evaluation of any new tools to determine their potential usefulness in the scope of JRA4. We believe that by providing specific and relevant performance data about each of the Grid components of the different middlewares as well as by making the tools available to perform such evaluations, JRA4 represents a significant contribution to the global Grid community. Through the development of an easy-to-use performance instrumentation API's for Grid components, JRA4 will facilitate performance evaluation and make optimization a natural part of the Grid software development process, ultimately resulting in an improved end-user experience: more work done per unit time.

5 References

- [1] Alexius, Elahi, Fiorentino, Hedman, Mucci, Netzer, *OMII Europe JRA4 Whitepaper*, May 2007. Internal working report.
- [2] Elahi, Mucci, Alexius, Hedman, *An Overview of the Approaches to Grid Benchmarking*, May 2007. In preparation.